



STIC Search Report

Biotech-Chem Library

STIC Database Tracking Number: 150258

TO: Ruixiang Li
Location: REM/4D75/4C70
Art Unit: 1646
Friday, April 15, 2005

Case Serial Number: 10/664667

From: Toby Port
Location: Biotech-Chem Library
REM1-A59
Phone: 272-2523

toby.port@uspto.gov

Search Notes

Dear Examiner Li,

Here are the results of your search.
Please feel free to contact me if you have any questions.

Toby Port

This Page Blank (uspto)

STIC-Biotech/ChemLib

150258

From: Li, Ruixiang
Sent: Friday, April 08, 2005 11:20 AM
To: STIC-Biotech/ChemLib
Subject: Sequence search of Application No.10/664,667

Please do a standard search on:

- (i). SEQ ID NOS: 4, 5, and 6 against commercial nucleic acid databases;
- (ii). SEQ ID NO: 5 against commercial amino acid databases.

Thank you very much!

Ruixiang Li
GAU 1646
REM 4D75
Mail Box 4C70
(571) 272-0875

RECEIVED
APR - 8 2005
STIC

STAFF USE ONLY

Searcher: _____
Searcher Phone: 2- _____
Date Searcher Picked up: _____
Date Completed: _____
Searcher Prep/Rev. Time: _____
Online Time: _____

Type of Search

NA#: _____ AA#: _____
Interference: _____ SPDI: _____
S/L: _____ Oligomer: _____
Encode/Transl: _____
Structure#: _____ Text: _____
Inventor: _____ Litigation: _____

Vendors and cost where applicable

STN: _____
DIALOG: _____
QUESTEL/ORBIS: _____
LEXIS/NEXIS: _____
SEQUENCE SYSTEM: _____
WWW/Internet: _____
Other(Specify): _____

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 14, 2005, 11:57:53 ; Search time 5363.18 Seconds
(without alignments)
17643.956 Million cell updates/sec

Title: US-10-664-667-4
Perfect score: 2486
Sequence: 1 taatcacgactacataagg.....agtgcaacccaattcaatt 2486

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 34239544 seqs, 19032134700 residues
Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : EST:*
1: gb_est1:*
2: gb_est2:*
3: gb_hic:*
4: gb_est3:*
5: gb_est4:*
6: gb_est5:*
7: gb_est6:*
8: gb_gss1:*
9: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2311.4	93.0	3509	3	BC038795 Homo sapi
2	1360.2	54.7	2822	3	AK052873 Mus muscu
3	1333.4	53.6	3583	3	AK085901 Mus muscu
4	747.8	30.1	771	5	BQ425502 AGENCOURT
5	590.2	23.7	647	2	AW970780 EST382863
6	539.8	21.7	982	2	BF159363 601770339
7	534.4	21.5	819	4	BG863804 602798701
8	520.2	20.9	599	1	AA392507 2530b10.x
9	516.2	20.8	812	4	BG916782 602816031
10	515	20.7	3855	3	AK040883 Mus muscu
11	507.4	20.4	509	2	AW009250 w879g09.x
12	486.4	19.6	690	2	BB636845 BB636845
13	477.4	19.2	482	4	BM659843 UI-E-DX1-
14	476	19.1	479	5	BM993280 UI-H-DT0-
15	474.2	19.1	633	1	AA524536 r945g03.s
16	468.4	18.8	471	5	BX281802 BX281802
17	468.8	18.7	680	7	CN258480 170005321
18	457.6	18.4	753	4	BG916963 602816251
19	455	18.3	558	1	AA131133 2016b04.x
20	451.2	18.1	468	1	AA443637 2w35f07.s
21	451	18.1	451	2	BE326280 hw01d11.x
22	439.4	17.7	441	1	AA716422 2964c06.s
23	424.4	17.1	869	2	BF158974 601766476
24	416.8	16.8	420	2	AW167811 xg56f07.x

C 25	399.4	16.1	401	1	AI989794	AI989794 w827c09.x
C 26	398.2	16.0	577	7	W93698	W93698 zd96c01.b1
C 27	398	16.0	398	1	AI204695	AI204695 qd44f06.x
C 28	391.4	15.7	393	1	AI206303	AI206303 qr26g08.x
C 29	385.2	15.5	461	2	BE149368	BE149368 RC1-HT025
C 30	382.2	15.4	472	1	AA815396	AA815396 ai61c08.s
C 31	378.8	15.2	395	2	AW102822	AW102822 xd38f11.x
C 32	362.4	14.6	2724	9	AY411734	AY411734 Mus muscu
C 33	359	14.4	368	1	AA195671	AA195671 zr32g12.r
C 34	357	14.4	381	2	AW104260	AW104260 xd70g09.x
C 35	355.6	14.3	2724	9	AY411732	AY411732 Homo sapi
C 36	349.8	14.1	443	1	AI131316	AI131316 tc18a07.x
C 37	344.6	13.9	373	1	AI686439	AI686439 tw99f07.x
C 38	322.4	13.0	640	7	CO351740	CO351740 DR_AOV NR
C 39	314.6	12.7	400	7	W93749	W93749 zd96c01.r1
C 40	310.8	12.5	377	1	AA235322	AA235322 zt30b10.s
C 41	310.2	12.5	467	1	AA131081	AA131081 zo16b04.s
C 42	282.4	11.4	310	7	Z44115	Z44115 HSC1SG111.n
C 43	277.2	11.2	1719	2	BF144122	BF144122 601788812
C 44	274	11.0	2593	9	AY411733	AY411733 Pan trogl
C 45	267.8	10.8	311	1	AA195672	AA195672 zr32g12.s

ALIGNMENTS

RESULT 1	BC038795	3509 bp	mRNA	linear	HTC 19-NOV-2003
LOCUS	BC038795				
DEFINITION	Homo sapiens leucine-rich repeat-containing G protein-coupled receptor 6, mRNA (cdna clone IMAGE:5220507), with apparent retained intron.				
ACCESSION	BC038795				
VERSION	BC038795.1	GI:24433474			
KEYWORDS	HTC.				
SOURCE	Homo sapiens				
ORGANISM	Homo sapiens				
REFERENCE	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.				
AUTHORS	1 (bases 1 to 3509) Strausberg, R.L., Feingold, E.A., Grouse, L.H., Derge, J.G., Klausner, R.D., Collins, F.S., Wagner, L., Shenmen, C.M., Schuler, G.D., Altschul, S.F., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Bhat, N.K., Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Hsieh, F., Diatchenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L., Stapleton, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L., Scheetz, T.E., Brownstein, M.J., Usdin, T.B., Toshiyuki, S., Carninci, P., Prange, C., Raha, S.S., Loquellano, N.A., Peters, G.J., Abramson, R.D., Mullaly, S.J., Bosak, S.A., McEwan, P.J., McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S., Villalon, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A., Fahey, J., Helton, E., Kettman, M., Madan, A., Rodriguez, S., Sanchez, A., Whiting, M., Madan, A., Young, A.C., Shevchenko, Y., Bouffard, G.G., Blakesley, R.W., Touchman, J.W., Green, E.D., Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M., Butlerfield, Y.S., Krzywicki, M.I., Skalek, U., Smalish, D.E., Scherch, A., Schein, J.E., Jones, S.J. and Marra, M.A.				
TITLE	Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences				
JOURNAL	Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)				
MEDLINE	22388257				
PUBMED	12477932				
REFERENCE	2 (bases 1 to 3509) Strausberg, R.				
AUTHORS	Direct Submission				
TITLE	Submitted (25-OCT-2002) National Institutes of Health, Mammalian				
JOURNAL	Gene Collection (MGC), Cancer Genomics Office, National Cancer				
	Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,				
	USA				
REMARK	NIH-MGC Project URL: http://mgc.nci.nih.gov				
COMMENT	Contact: MGC help desk				
	Email: cgapbs-remail.nih.gov				

Tissue Procurement: Life Technologies, Inc.
 cDNA Library Preparation: Life Technologies, Inc.
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: National Institutes of Health Intramural
 Sequencing Center (NISC),
 Gaithersburg, Maryland;
 Web site: <http://www.nisec.nih.gov/>
 Contact: nisc_mgc@hgrl.nih.gov
 Akhter, N., Ayele, K., Beckstrom-Sternberg, S.M., Benjamin, B.,
 Blakesley, R.W., Bouffard, G.G., Breen, K., Brinkley, C., Brooks, S.,
 Dietrich, N.L., Granito, S., Guan, X., Gupta, J., Haghighi, P.,
 Hansen, N., Ho, S.-L., Karlins, E., Kwong, P., Laric, P., Legaspi, R.,
 Maduro, Q.L., Masliello, C., Maskeri, B., Mastrian, S.D., McCluskey, J.C.,
 McDowell, J., Pearson, R., Stantripop, S., Thomas, P.J., Touchman, J.W.,
 Tsurgeon, C., Vogt, J.L., Walker, M.A., Wetherby, K.D., Wiggins, L.,
 Young, A., Zhang, L.-H. and Green, E.D.

Clone distribution: MGC clone distribution information can be found
 through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
 Series: IRAK Plate: 64 Row: n Column: 17
 This clone was selected for full length sequencing because it
 passed the following selection criteria: Hexamer frequency ORF
 analysis

This clone has the following problem: retained intron.

FEATURES

Location/Qualifiers
 1..3509
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:5220507"
 /tissue type="Pancreas, Spleen, adult pooled"
 /clone_lib="NTH MGC_120"
 /lab_host="DH10B"
 /note="Vector: pCMV-SPORT6"

ORIGIN

Query Match 93.0%; Score 2311.4; DB 3; Length 3509;
 Best Local Similarity 99.7%; Pred. No. 0;
 Matches 2326; Conservative 0; Mismatches 5; Indels 1; Gaps 1;

Qy 71 GCGTCCGTGGAGCGGAGCAGGGTCTGAGCTGCGGCTCATCCAGCCTCTCTTGTGCCC 130
 Db 1063 GGGGCGATGGAGCGGAGCAGGGTCTGAGCTGCGGCTCATCCAGCCTCTCTTGTGCCC 1122

Qy 131 CTAGGGGCTCCACACACGCGCATCTGGGAAATGGAGCTNGACACCTTCAGCCAGCTG 190
 Db 1123 CTAGGGGCTCCACACACGCGCATCTGGGAAATGGAGCT-GACACCTTCAGCCAGCTG 1181

Qy 191 AGCTCCCTGCAAGCCCTGGATCTTAGCTGGAAAGCGCATCCGGTCCATCCACCCTGAGGCC 250
 Db 1182 AGCTCCCTGCAAGCCCTGGATCTTAGCTGGAAAGCGCATCCGGTCCATCCACCCTGAGGCC 1241

Qy 251 TTCTCCACCTTGCACTCCCTGGTCAAGCTGGACCTGACAGACAACAGCTGACCACTG 310
 Db 1242 TTCTCCACCTTGCACTCCCTGGTCAAGCTGGACCTGACAGACAACAGCTGACCACTG 1301

Qy 311 CCCCTGGCTGCACTTGGGGCTTGATGCACTGAAGCTCAAGGGAACCTTGCTCTTCC 370
 Db 1302 CCCCTGGCTGCACTTGGGGCTTGATGCACTGAAGCTCAAGGGAACCTTGCTCTTCC 1361

Qy 371 CAGGCTTCTCCAAAGGACAGTTTCCCAAACTGAGGATCCTGGAGGTGCCCTTATGCTTAC 430
 Db 1362 CAGGCTTCTCCAAAGGACAGTTTCCCAAACTGAGGATCCTGGAGGTGCCCTTATGCTTAC 1421

Qy 431 CAGTGTCTTCCCTATGGGATGTGCCAGCTTCTTCAAGGCTCTGGGCACTGGGAGGCT 490
 Db 1422 CAGTGTCTTCCCTATGGGATGTGCCAGCTTCTTCAAGGCTCTGGGCACTGGGAGGCT 1481

Qy 491 GAAGACCTTCACCTTGATGAGGAGTCTTCAAAAGGCCCTGGGCTCTTGGCCAGA 550
 Db 1482 GAAGACCTTCACCTTGATGAGGAGTCTTCAAAAGGCCCTGGGCTCTTGGCCAGA 1541

Qy 551 CAAGCAGAGAACCACTATGACAGGACCTGGATGAGCTCCAGCTGGAGATGGAGGACTCA 610

Db 1542 CAAGCAGAGAACCACTATGACAGGACCTGGATGAGCTCCAGCTCCAGCTGGAGATGGAGGACTCA 1601

Qy 611 AAGCCACACCCAGTGTCCAGTGTAGCCCTACTCCAGGCCCTTTCAGGCCCTGTGAGTAC 670

Db 1602 AAGCCACACCCAGTGTCCAGTGTAGCCCTACTCCAGGCCCTTTCAGGCCCTGTGAGTAC 1661

Qy 671 CTCTTTTAAAGCTGGGCGATCCCGCTGCGGTGTGGGCGCATCGGTGTGTCTCTCGTGTCTC 730

Db 1662 CTCTTTTAAAGCTGGGCGATCCCGCTGCGGTGTGGGCGCATCGGTGTGTCTCTCGTGTCTC 1721

Qy 731 TGCAATGGACTGTGTCTGTGACCGTGTTCGTGTGGCGGCTGTGCCCTGTGCCCTCCGGTGC 790

Db 1722 TGCAATGGACTGTGTCTGTGACCGTGTTCGTGTGGCGGCTGTGCCCTGTGCCCTCCGGTGC 1781

Qy 791 AAGTTTGTGTAGTGTGGATTGCAAGGCGCAACACCTTGACTGCGCATTTCTGTGGGCTT 850

Db 1782 AAGTTTGTGTAGTGTGGATTGCAAGGCGCAACACCTTGACTGCGCATTTCTGTGGGCTT 1841

Qy 851 CTAGCCTCAGTCGATGCCCTGACCTTTTGTGTCAGTTCTCTGAGTAGCGAGCCGCTGGGAG 910

Db 1842 CTAGCCTCAGTCGATGCCCTGACCTTTTGTGTCAGTTCTCTGAGTAGCGAGCCGCTGGGAG 1901

Qy 911 ACGGGGCTAGGCTGCCGGGCCACTGGCTTCTGTGCACTACTTGGGTGGAGGCAATCGGTG 970

Db 1902 ACGGGGCTAGGCTGCCGGGCCACTGGCTTCTGTGCACTACTTGGGTGGAGGCAATCGGTG 1961

Qy 971 CTGCTGTCTACTCTGGGCGCAGTGCAGGTCTCCGTCTCTCTGTGTCCGGGCTTAT 1030

Db 1962 CTGCTGTCTACTCTGGGCGCAGTGCAGGTCTCCGTCTCTCTGTGTCCGGGCTTAT 2021

Qy 1031 GGGAAAGTCCCTCCCTGGGCGAGGTTTCGAGAGGGGCTCTAGGCTGCTGGCACTGGCA 1090

Db 2022 GGGAAAGTCCCTCCCTGGGCGAGGTTTCGAGAGGGGCTCTAGGCTGCTGGCACTGGCA 2081

Qy 1091 GGGTGGCCCGCAGCTGCCCTTGGCTCAGTGGGAGAAATACGGGGCTCCCTCCACTCTGC 1150

Db 2082 GGGTGGCCCGCAGCTGCCCTTGGCTCAGTGGGAGAAATACGGGGCTCCCTCCACTCTGC 2141

Qy 1151 CTGCCCTACGGGCGACCTGAGGCTCAGCAGCAGCCCTGGGCTTCAGGTTGGGCTCGGTG 1210

Db 2142 CTGCCCTACGGGCGACCTGAGGCTCAGCAGCAGCCCTGGGCTTCAGCGTGGGCTCGGTG 2201

Qy 1211 ATGATGAATCTCTCTCTTCTTCTGCTGCGCGGTGCTTACATCAAACTGTACTGTGAC 1270

Db 2202 ATGATGAATCTCTCTCTTCTTCTGCTGCGCGGTGCTTACATCAAACTGTACTGTGAC 2261

Qy 1271 CTGCCGCGGGCGACTTTGAGGCGGTGTGGGACTGCGCCATGTTGAGGCACTGGGCTGG 1330

Db 2262 CTGCCGCGGGCGACTTTGAGGCGGTGTGGGACTGCGCCATGTTGAGGCACTGGGCTGG 2321

Qy 1331 CTGATCTTCGAGACGGGCTCTCTACTGTCCGTGGGCTTCTCTCAGCTTGGCTCCATG 1390

Db 2322 CTGATCTTCGAGACGGGCTCTCTACTGTCCGTGGGCTTCTCTCAGCTTGGCTCCATG 2381

Qy 1391 CTGGGCTCTTCTCCCTGTACGCGGAGCGTCAAGTCTGTCTGTGGTGGTGTCTGCC 1450

Db 2382 CTGGGCTCTTCTCCCTGTACGCGGAGCGTCAAGTCTGTCTGTGGTGGTGTCTGCC 2441

Qy 1451 CTGCTCGCTCGCTCAACCCCACTGCTGTCTTCAACCCCACTTTCGGGATGAC 1510

Db 2442 CTGCTCGCTCGCTCAACCCCACTGCTGTCTTCAACCCCACTTTCGGGATGAC 2501

Qy 1511 CTTGCGGGCTTTCGGGCGCGGAGGAGCTCAGGGGCCCTTAGGCTATGCTGGGCGGG 1570

Db 2502 CTTGCGGGCTTTCGGGCGCGGAGGAGCTCAGGGGCCCTTAGGCTATGCTGGGCGGG 2561

Qy 1571 GAGCTGAGAGAGAGCTCTCTGTATTCTACCCAGGCCCTGGTGGGCTTCTCTGTATGTGAT 1630

Db 2562 GAGCTGAGAGAGAGCTCTCTGTATTCTACCCAGGCCCTGGTGGGCTTCTCTGTATGTGAT 2621

Qy 1631 CTCAATCTGGAAGCTTCTGAAGCTGGGCGGCCCTCTGGGCTGGAGACTTATGGCTTCCCC 1690

Db 2622 CTCATCTGGGAAGCTTCTGAAGCTTGGGCGGCCCCCTGGGCTGGAGACCTATGCTTCCTCCC 2681

Qy 1691 TCAGTGACCCCTCATCTCTCTGTGTCAGCAGCAGCGGGCCCCCAGGCTCGAGGGCAGCCATTGT 1750

Db 2682 TCAGTGACCCCTCATCTCTGTGTCAGCAGCAGCGGGCCCCCAGGCTCGAGGGCAGCCATTGT 2741

Qy 1751 GTAGAGCCAGAGGGGAACACATTTTGGGAACCCCCAAACCCCTCATGGATGGAGAACTGCTG 1810

Db 2742 GTAGAGCCAGAGGGGAACACATTTTGGGAACCCCCAAACCCCTCATGGATGGAGAACTGCTG 2801

Qy 1811 CTGAGGCGAGAGGGATCTACGCCAGCAGCTGGAGGCTTGTACAGGGGGTGGCGGCTTTCAG 1870

Db 2802 CTGAGGCGAGAGGGATCTACGCCAGCAGCTGGAGGCTTGTACAGGGGGTGGCGGCTTTCAG 2861

Qy 1871 CCCTCTGGCTGGCCCTTTGCTTTCACACGCTGTAATATCCCTCCCACTTCTCTCTTCCTCCC 1930

Db 2862 CCCTCTGGCTGGCCCTTTGCTTTCACACGCTGTAATATCCCTCCCACTTCTCTCTTCCTCCC 2921

Qy 1931 TCTCTTCCCTTTCCTCTCTCCCTCGGTGAATGATGGCTGCTTCTTCTTCTTCTTCTTCTTCAAC 1990

Db 2922 TCTCTTCCCTTTCCTCTCTCCCTCGGTGAATGATGGCTGCTTCTTCTTCTTCTTCTTCTTCAAC 2981

Qy 1991 CAAAACTCAGCAGCTGTGATCTATAGCAGATGGCCCGAGTACCTGCTGCTCCACTGATCACCT 2050

Db 2982 CAAAACTCAGCAGCTGTGATCTATAGCAGATGGCCCGAGTACCTGCTGCTCCACTGATCACCT 3041

Qy 2051 CTCTCTGTGACCATCACCACGGGTGCTCTTGGCCTGGCTTTCCCTTGGCTTCCCTCA 2110

Db 3042 CTCTCTGTGACCATCACCACGGGTGCTCTTGGCCTGGCTTTCCCTTGGCTTCCCTCA 3101

Qy 2111 GCTTCACCTTGATCTGGGCTCTTCTTGTCTGATCTCTGAAGCTGTGACGACGACCTG 2170

Db 3102 GCTTCACCTTGATCTGGGCTCTTCTTGTCTGATCTCTGAAGCTGTGACGACGACCTG 3161

Qy 2171 GACTTTTCTCTTAAAGGAATCAGGAGTAAAGCAGTGAAGGGTGGAGGGTGA 2230

Db 3162 GACTTTTCTCTTAAAGGAATCAGGAGTAAAGCAGTGAAGGGTGGAGGGTGA 3221

Qy 2231 TCAGGGCAGCTGGAGGACCTCACAGAAAGGCTTGAAGGTGATTTCCCGT 2290

Db 3222 TCAGGGCAGCTGGAGGACCTCACAGAAAGGCTTGAAGGTGATTTCCCGT 3281

Qy 2291 GACTCATGATAGGATACAAATGTGTTCCATGTACCTTAATCTTGACATATGCCATGC 2350

Db 3282 GACTCATGATAGGATACAAATGTGTTCCATGTACCTTAATCTTGACATATGCCATGC 3341

Qy 2351 ATAAGACTTCTTATAAATAGCTTTGGAGAGATTAATAAAAAAAAAAAAA 2403

Db 3342 ATAAGACTTCTTATAAATAGCTTTGGAGAGATTAATAAAAAAAAAAAAA 3394

RESULT 2

AK052873 AK052873 2822 bp mRNA linear HTC 03-APR-2004

LOCUS Mus musculus 16 days neonate heart cDNA, RIKEN full-length enriched

DEFINITION library, clone:D830015D13 product:CDNA FLJ14471 FIS, CLONE

MAMMA1001030, WEAKLY SIMILAR TO LUTROPIN-CHORIOGNADOTROPIC HORMONE

RECEPTOR homolog [Homo sapiens], full insert sequence.

ACCESSION AK052873

VERSION AK052873.1 GI:26095426

KEYWORDS HTC; CAP trapper.

SOURCE Mus musculus (house mouse)

ORGANISM Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

1

REFERENCE Carninci, P. and Hayashizaki, Y.

AUTHORS High-efficiency full-length cDNA cloning

TITLE Meth. Enzymol. 303, 19-44 (1999)

JOURNAL 99279253

MEDLINE 10349636

PUBMED

2

REFERENCE Carninci, P., Shibata, Y., Hayateu, N., Sugahara, Y., Shibata, K.,

Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
Normalization and subtraction of cap-trapper-selected cDNAs to
prepare full-length cDNA libraries for rapid discovery of new genes
Genome Res. 10 (10), 1617-1630 (2000)
20499374
11042159

3
Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P.,
Konno, H., Akiyama, J., Nishi, K., Kitagawa, T., Tashiro, H., Itoh, M.,
Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A.,
Yanamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kaehiwagi, K.,
Fujiwaka, S., Inoue, K., Izawa, M., Ohara, S., Watahiki, M.,
Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsuura, S., Kawai, J.,
Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.
RIKEN integrated sequence analysis (RISA) system--384-format
sequencing pipeline with 384 multicapillary sequencer
Genome Res. 10 (11), 1757-1771 (2000)
20530913
11076861

4
The RIKEN Genome Exploration Research Group Phase II Team and the
FANTOM Consortium.
Functional annotation of a full-length mouse cDNA collection
Nature 409, 685-690 (2001)
5

5
The FANTOM Consortium and the RIKEN Genome Exploration Research
Group Phase I & II Team.
Analysis of the mouse transcriptome based on functional annotation
of 60,770 full-length cDNAs
Nature 420, 563-573 (2002)
6 (Bases 1 to 2822)

Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Bono, H., Carninci, P.,
Fukuda, S., Furuno, M., Hanagaki, T., Hara, A., Hashizume, W.,
Hayashida, K., Hayateu, N., Hiramoto, K., Hiraoka, T., Hirozane, T.,
Hori, F., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kasukawa, T.,
Kato, H., Kawai, J., Kojima, Y., Kondo, S., Konno, H., Kouda, M.,
Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M.,
Nakamura, M., Nishi, K., Nomura, K., Numasaki, R., Ohno, M., Ohsato, N.,
Okazaki, Y., Saito, R., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N.,
Sano, H., Sasaki, D., Shibata, K., Shinagawa, A., Shiraki, T.,
Sogabe, Y., Tagami, M., Tagawa, A., Takahashi, F., Takaku-Akahira, S.,
Takeda, Y., Tanaka, T., Tomaru, A., Toyota, T., Yasunishi, A.,
Muramatsu, M. and Hayashizaki, Y.
Direct Submission

Submitted (16-JUL-2001) Yoshihide Hayashizaki, The Institute of
Physical and Chemical Research (RIKEN) Laboratory for Genome
Exploration Research Group, RIKEN Genomic Sciences Center (GSC),
RIKEN Yokohama Institute, 1-7-22 Suenho-cho, Tsurumi-ku, Yokohama,
Kanagawa 230-0045, Japan (E-mail: genome-res@gscc.riken.jp,
URL: http://genome.gsc.riken.jp/, Tel: 81-45-503-9222,
Fax: 81-45-503-9216)

cDNA library was prepared and sequenced in Mouse Genome
Encyclopedia Project of Genome Exploration Research Group in Riken
Genomic Sciences Center and Genome Science Laboratory in RIKEN.
Division of Experimental Animal Research in Riken contributed to
prepare mouse tissues.

Please visit our web site for further details.
URL: http://genome.gsc.riken.jp/
URL: http://fantom.gsc.riken.jp/.

Location/Qualifiers
1. .2822
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="FANTOM_DB:D830015D13"
/db_xref="taxon:10090"
/clone="D830015D13"
/tissue type="heart"
/clone_lib="RIKEN full-length enriched mouse cDNA library"
/dev_stage="16 days neonate"
1. .2822

misc_feature
/note="CDNA FLJ14471 FIS, CLONE MAMMA1001030, WEAKLY
SIMILAR TO LUTROPIN-CHORIOGNADOTROPIC HORMONE RECEPTOR

homolog [Homo sapiens] (SPTR|Q96K69, evidence: PASTY,
85.8%ID, 92.1%length, match=637)"

ORIGIN							
Query Match							
Best Local Similarity 54.7%; Score 1360.2; DB 3; Length 2822;							
Matches 1775; Conservative 0; Mismatches 439; Indels 72; Gaps 8;							
QY	136	GGCTCCCAACACACCGCATCTGGGAATTTGGAGTNGACACCTTCACCCAGCTGAGCTC	195	Db	1433	GGCCGACGACCTGCGCTGGGCTCGTGGGAGAGTATGGCGCCTCCCACTCTGCCTGCC	1492
Db	474	GGCTCCCGACATAAACCGGATCAAGGAAATTTGGTGC-AGATACCTTCAGCCAGCTGGGCTC	532	QY	1156	CTACGGCCACCTGAGGGTCAGCCAGACGCCCTGGGCTTACCCGTGGCCCTGGTGATGAT	1215
QY	196	CCTGCAAGCCCTGGATCTTTAGCTGGAACGCCATCCGGTCCATCCACCCCTGAGGCCCTTCTC	255	Db	1493	CTACGCCCCACCCGAGGGCCGGCTGGCCCTGGGCTTCGCTGTAGCCCTGGTGATGAT	1552
Db	533	CTTGAAGCTTTAGACCTGAGTTGGAATGCCATCCGTGCCATCCACCTGAGGCTTTCTC	592	QY	1216	GAACTCTCTCTGTTTCTGTCGTCGGCGGTCCTACATCAAACTGTACTGTGACCTGCC	1275
QY	256	CACCTGCACTCCCTGGTCAAGCTGGACCTGACAGACAAACAGCTGACCACTGCCCT	315	Db	1553	GAACTCGCTCTGCTTCTTGGTGGCGGGCCCTACATCAAGCTCTACTGTGACCTGCC	1612
Db	593	AACCCCTCGATCTTGGTTAAGCTGGACCTGACTGACAAACAGCTGACCACTGCCCT	652	QY	1276	GCGGGGCGACTTTGAGGCCGTGTGGGACTGCGCCATGGTGTGAGGCACGT-GGCCTGGCTCA	1334
QY	316	GGCTGGACTTGGGGCTTGATGCATCTGNAAGCTCAAGGGACCTTGCTCTCTCCAGGC	375	Db	1613	ACGGGGTGACTTTGAGGCCGTGTGGGACTGCGCCATGGTGGCCCTGGCTCA	1672
Db	653	GGCTGGGCTGGAGGCCCTGATGCACCTGAAGCTCAAGGGAACTTGGGCCCTGTCTCAGGC	712	QY	1335	TCCTTGGAGAGCGGCTCTCTACTGTCCCGTGGGCTTCTCAGCTTCCGCTCCATGTCTGG	1394
QY	376	CTTCTCCAAGACAGTTTCCCAAACTGAGGATCTCGAGGTGCTTATGCTTACCAAGTG	435	Db	1673	TCCTTGGAGATGGCTCTCTACTGTGCCCGTGGGCTTCTCAGCTTTCGCTCTATGTCTGG	1732
Db	713	CTTCTCCAAGACAGTTTCCCAAACTGAGGATCTCGAGGTGCTTATGCTTACCAAGTG	772	QY	1395	GCCTCTTCCCTGTCAACGCCGAGCCGTCAAGTCTGCTGCTGCTGGTGGTGGTGGCCCTGC	1454
QY	436	CTGTCCCTATGGGATGTGTGCAGCTTTCAAGGCCCTTGGGCAAGTGGGAGGCTGAAGA	495	Db	1733	GCCTCTTCCCTGTCAACGCCGAGCCGTCAAGTCTGCTTCTGCTGGTGGTGGTGGCCCTGC	1792
Db	773	CTGTCCCTATGGGATGTGTGCAGCTTTCAAGGCCCTTGGGCAAGTGGGAGGCTGAAGA	832	QY	1455	CTGCTTCCCTCAACCCACTGCTGTACTGTCTTTCAACCCCACTTCCGGGATGACCTTC	1514
QY	496	CCTTCACTTGTATGATGAGGAGTCTTCAAAAGGCCCTGGGCTTCTTGGCAGACAGC	555	Db	1793	CTGCTTCCCTCAACCCACTGCTGTACTGTCTTTCAACCCCACTTCCGGGATGACCTTC	1852
Db	833	CTTTTCATCCAGAGGAAGAGGAGCCAAAGAGGCCCTTGGGCTCTCTTGTGTGACAAAGC	892	QY	1515	GGCGGCTTGGGCCCGGCGAGGGGACTCAGGGCCCTTAGGCTATGCTGGCGCCGGGAGC	1574
QY	556	AGAGAACCACTATGACAGGACCTGGATGAGCTCCAGCTGGAGATGGAGGACTCAAGCC	615	Db	1853	GGCGGCTTGGGCCCGGCGAGGGGCTCAGGGCCCTTAGGCTATGCTGGCGCCGGGAGC	1912
Db	893	TGAGAACCACTATGACCTTACCTTGGATGAGCTCCAGATGGGACAGAGGACTCAAGCC	952	QY	1575	TGAGAGAGAGCTCTGCTGATTTACACGAGCCCTGGTGTAGCTTCTCTGATGTGATCTCA	1634
QY	616	ACACCCAGTGTCCAGTGTAGCCCTTATCCAGGCCCTTCAAGGCCCTGTGAGTACCTTT	675	Db	1913	TGAGAGAGAGCTCTGCTGATTTACACGAGCCCTGGTGTAGCTTCTCTGATGTGATCTCA	1972
Db	953	ACACCCAGTGTCCAGTGTAGCCCTTATCCAGGCCCTTCAAGGCCCTGTGAGTACCTTT	1012	QY	1635	TTCTGGAAGCTTTCTGGAAGCTGGCGGCCCTTGGGCTGGAGACCTATGGCTTCCCTCAG	1694
QY	676	TGAAGCTTGGGGCATCCGCTGGCGTGGGCACTGCTGTCTCCGCTCTGCA	735	Db	1973	TTCTGGAAGCTTTCTGGAAGCTGGCGGCCCTTGGGCTGGAGACCTATGGCTTCCCTCAG	2032
Db	1013	TGAAGCTTGGGGCATCCGCTTGTGTGTGGGCACTGCTGTACTCTCCGCTCTGTAA	1072	QY	1695	TGACCTCTATCTCTGTCTGACAGCCAGGGGCCCGGCTGGAGGGAGCCATTTGTGTAG	1754
QY	736	TGGACTGTGTGCTGACCGCTTGTGCTGGCGGGCTTGGCCCTTGGCCCGGCTCAAGTT	795	Db	2033	TGACCTCTATCTCTGTCTGACAGCCAGGGGCCCGGCTGGAGGGAGAAACATTTGTGTAG	2092
Db	1073	CGGGCTGTGTGCTGACAGCTTTTGGCAGCGGACCCAGCCCGCTGTCTCCCGCTCAAGCT	1132	QY	1755	AGCCAGAGGGGAAACCACTTTGGGAAACCCCAACCTCCATGTGATGGAGAACTGCTGTGA	1814
QY	796	TGTGGTAGGTGCGATTTGAGCGCCAAACCTTGA CTGGGCAATTTCTGTGGCCCTTCTAGC	855	Db	2093	AGTCTGATGGAACCAAGTTTGGGAAACCCCAACCTCCATGAAAGGGAGAACTGCTGTGA	2152
Db	1133	TGTGGTGGTGGATGGCAGCGCCAAAGCCCTGTGCGGGCAATTTCTGTGTCTCTCTGGC	1192	QY	1815	GGGACAGAGGATCTTACGCCACAGGCTGGAGGCTTGTGAGGGGGTGGCGGCTTTTCAGGCCCT	1874
QY	856	CTCAGTCGATCCCTGACCTTTGTTCAGTTCTCTGAGTACGAGGCCCGCTGGGAGACGG	915	Db	2153	AGGACAGAGGAGCCACTTTTGGCAGGCTGTGGCTCTTCCGTGGGTGGAGCCCTCTGGGCCCT	2212
Db	1193	CTCGGTGACCCCTTGACCTATGTGTCAGTTTCGCTGAGTATGAGCCCGCTGGGAGACGG	1252	QY	1875	CTGGCTTGGCTTGTGCTTCAAGTGAATAATCCCTCCCACTTCTTCTCTTCCCTCTCTC	1934
QY	916	GCTAGGCTGCGGGCCACTGCTTCTCGGCACTTCTGGGCTGGAGGATCGGTGCTGCT	975	Db	2213	CTGGCTCTCTCTTGTGCTCTCACTTGAATAATCCCTCTCTGTTGTGTCTCTTCCCGTCC	2272
Db	1253	TCTGGCTTGCAGGATACGGGCTTCTGGCTGTCTGGTTTCAGAGGGTGGGTGCTGCT	1312	QY	1935	TTCCCTTCTCTCTCTCCCTCTGGTGAATGATGGCTTCTTAAACAAATACAAACAA	1994
QY	976	GCTCACTGTGCCCGAGTGCAGTGCAGGTTCTCCGCTCTCTGTGTGTGGGCGCTATGGGAA	1035	Db	2273	-----AATGATGGCTGCTTATATAAAGAAAGAACATCTCA	2306
Db	1313	GCTCACACTGGCGCGGTGACAGTGCAGCATCTCCGTGACCTGCGTCCGAGCCTACGGGAA	1372	QY	1995	ACTCAGCAGTGTGATCTATAGCAGGATGGCCCACTAGTCTGGCTCCACTGATCACTCTCT	2054
QY	1036	GTCCCTTCTTGGGACAGCTTCAGACAGGGGTCTTAGGCTGCTTGGCACTGGGAGGGCT	1095	Db	2307	ACTCC-----ATAGCAAGATGGCCCAACACTCTGACTCCATTGTT--CTCTCT	2352
Db	1373	GGCGCGCTGCTGGCAGCGTCCGCGCAGCGCACTGGGATGCTTGGCGCTGGCGGGCT	1432	QY	2055	CTGTGTGACCAATACCAACGGGTGCTCTTGGCTGGCTTTCCTTGGCCCTTCTCAGCTT	2114
QY	1096	GGCGCGCAGCTGCGCCCTGGGCTCAGTGGGAGAAATACGGGGCCCTCCCACTCTGCCTGCC	1155	Db	2353	CCACGACCCCTAAACCAATGAGTGTCCAAAGTCTTGTGTTGTCTTGTGSCCT---TCAGCTT	2409

Db 2470 TGCATAGGAAAGGAGAAACCAAGACAGTGAAGGTTATGGGGCCCTGACAGAGCCA 2529
QY 2235 GGCACAGT-----GGACAGGGAGACCTTCACAGAGAAAGGCTCGAAGGTGATTTC 2284
Db 2530 TGATCAGTAAGTCAGAGTGTGGGGAGGTCTTCACAGAGCATGACACTGGAGACAACTA 2589
QY 2285 CGGTGACTCATGG-----ATAGGATACAAAATGTTGTTCCATGTACC 2327
Db 2590 CCAAGACATTGGAGAGTCTCCCTGTGACATATAGAAATATATAATGTTGTTGTTCC 2649
QY 2328 ATTAATCTTGACATATGTCATGATGATGATGATGATGATGATGATGATGATGATGAT 2387
Db 2650 ATTAATCTTGACATATGTCATGATGATGATGATGATGATGATGATGATGATGATGAT 2709
QY 2388 TAAAAA 2393
Db 2710 TGCATA 2715

RESULT 3
LOCUS AK085901
DEFINITION Mus musculus 3583 bp mRNA linear HTC 03-APR-2004
library, clone:D830026M09 product:CDNA FLJ14471 FIS, CLONE
MAMMA1001030, WEAKLY SIMILAR TO LUTROPIN-CHORIOGONADOTROPIC HORMONE
RECEPTOR homolog [Homo sapiens], full insert sequence.
ACCESSION AK085901
VERSION AK085901.1 GI:26103061
KEYWORDS HTC; CAP trapper.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus

REFERENCE 1
AUTHORS Carninci, P. and Hayashizaki, Y.
TITLE High-efficiency full-length cDNA cloning
JOURNAL Meth. Enzymol. 303, 19-44 (1999)
MEDLINE 99279253
PUBMED 10349636

REFERENCE 2
AUTHORS Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,
Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
TITLE Normalization and subtraction of cap-trapper-selected cDNAs to
prepare full-length cDNA libraries for rapid discovery of new genes
JOURNAL Genome Res. 10 (10), 1617-1630 (2000)
MEDLINE 20499374
PUBMED 11042159

REFERENCE 3
AUTHORS Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P.,
Konno, H., Akiyama, J., Nishi, K., Kitsunai, T., Tashiro, H., Itoh, M.,
Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A.,
Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kaehiwagi, K.,
Fujiwara, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watahiki, M.,
Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsura, S., Kawai, J.,
Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.
TITLE RIKEN integrated sequence analysis (RISA) system--384-format
sequencing pipeline with 384 multicapillary sequencer
JOURNAL Genome Res. 10 (11), 1757-1771 (2000)
MEDLINE 20530913
PUBMED 11076861

REFERENCE 4
AUTHORS The RIKEN Genome Exploration Research Group Phase II Team and the
FANTOM Consortium.
TITLE Functional annotation of a full-length mouse cDNA collection
JOURNAL Nature 409, 685-690 (2001)
REFERENCE 5
AUTHORS The FANTOM Consortium and the RIKEN Genome Exploration Research
Group Phase I & II Team.
TITLE Analysis of the mouse transcriptome based on functional annotation
of 60,770 full-length cDNAs
JOURNAL Nature 420, 563-573 (2002)
REFERENCE 6
AUTHORS Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Bono, H., Carninci, P.,

Fukuda, S., Furuno, M., Hanagaki, T., Hara, A., Hashizume, W.,
Hayashida, K., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hirozane, T.,
Hori, F., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kasukawa, T.,
Kato, H., Kawai, J., Kojima, Y., Kondo, S., Konno, H., Kouda, M.,
Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M.,
Nakamura, M., Nishi, K., Nomura, K., Numazaki, R., Ono, M., Ohsato, N.,
Okazaki, Y., Saito, R., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N.,
Sano, H., Sasaki, D., Shibata, K., Shinagawa, A., Shiraki, T.,
Sogabe, Y., Tagami, M., Tagawa, A., Takahashi, F., Takaku-Akahira, S.,
Takeda, Y., Tanaka, T., Tomaru, A., Toya, T., Yasunishi, A.,
Muramatsu, M. and Hayashizaki, Y.
Direct Submission
Submitted (16-APR-2002) Yoshihide Hayashizaki, The Institute of
Physical and Chemical Research (RIKEN), Laboratory for Genome
Exploration Research Group, RIKEN Genomic Sciences Center (GSC),
RIKEN Yokohama Institute, 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama,
Kanagawa 230-0045, Japan (E-mail: genome-res@gsc.riken.jp,
URL: http://genome.gsc.riken.jp/, Tel: 81-45-503-9222,
Fax: 81-45-503-9216)
COMMENT cDNA library was prepared and sequenced in Mouse Genome
Encyclopedia Project of Genome Exploration Research Group in Riken
Genomic Sciences Center and Genome Science Laboratory in RIKEN.
Division of Experimental Animal Research in Riken contributed to
prepare mouse tissues.
Please visit our web site for further details.
URL: http://genome.gsc.riken.jp/
URL: http://fantom.gsc.riken.jp/
FEATURES
Location/Qualifiers
source
1..3583
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="PANTOM:DB:D830026M09"
/db_xref="taxon:10090"
/clone="D830026M09"
/tissue_type="heart"
/clone_lib="RIKEN full-length enriched mouse cDNA library"
/dev_stage="16 days neonate"
misc_feature
95..3002
/note="CDNA FLJ14471 FIS, CLONE MAMMA1001030, WEAKLY
SIMILAR TO LUTROPIN-CHORIOGONADOTROPIC HORMONE RECEPTOR
homolog [Homo sapiens] (SPR:Q96K69, evidence: PASTY,
85.8%ID, 92.1%length, match=637)
putative"
polyA_signal
3563..3568
/note="putative"
polyA_site
3583
/note="putative"
ORIGIN
Query Match 53.6%; Score 1333.4; DB 3; Length 3583;
Best Local Similarity 77.4%; Pred. No. 0;
Matches 1772; Conservative 0; Mismatches 442; Indels 74; Gaps 10;
QY 136 GGCTCCCAACACACCCGATCTGGAAATTCGAGCTGACACCTTCAGCCAGCTGAGTTC 195
Db 1233 GGCTCCGAGTACACCCGATCAAGGAAATTCGAGCTGACACCTTCAGCCAGCTGAGTTC 1291
QY 196 CTGCAAGCCCTGATCTTAGCTGGAAGCCATCGGTCCATCCACCTGAGCCCTTCTC 255
Db 1292 CTTCGAAGCTTTAGACCTGAGTTCGAAATTCGATCCATCCACCTGAGGCTTCTC 1351
QY 256 CACCTGCACTCCCTGGTCAAGCTGGACCTCAGACAAACAGCTGACACACTGCCCTC 315
Db 1352 AACCTTCGATCTCTTGGTTAAGCTGGACCTGACCTGACACACAGCTGACACACTGCCCT 1411
QY 316 GGCTGGATTCGGGGCTTGATGATCTCTGAAGCTCAAAAGGAACTTGTCTCTCCAGGC 375
Db 1412 GGCTGGGCTGGAGGCTGATGACCTGACCTCAAGGAACTTGGGCTCTGTCTCAGGC 1471
QY 376 CTTCTCCAGACAGTTTCCCAAACTGAGATCTCTGAGGTGCTTATGCTTACCACTG 435
Db 1472 CTTCTCCAGACAGTTTCCCAAACTGAGATCTCTGAGGTGCTTACCACTGCTTACCACTG 1531

QY 436 CTGTCCTATGGATGTGTGCGAGCTTCTTCAAGGCTCTGGGAGTGGAGGCTGAAGA 495
 Db 1532 CTGTGCTTACGGCATCTGTGCGAGCTTCTTCAAGACCTTGGGAGTGGAGGCGGAGA 1591
 QY 496 CTTTCACTTGTATGATGAGGAGTCTTCAAAAAGGCGCTTGGGCTCTTGGCAGACAAGC 555
 Db 1592 CTTTTCATCCAGAGGAAGGAGGACCAAAAGAGGCGCTTGGTCTCTTGTGGACAAGC 1651
 QY 556 AGAGAACCATATGACCAAGGACCTGGATGAGTCCAGCTGGAGATGGAGACTCAAGGCC 615
 Db 1652 TGAGAACCATATGACCTGATGAGCTTGGATGAGTCCAGATGGGGAAGAGGACTCAAGGCC 1711
 QY 616 ACACCCCAAGTGTCCAGTGTAGCCCTACTCCAGGCGCCCTTCAA-GCCCTGTGAGTACCTCT 674
 Db 1712 ACACCCCAAGTGTCCAGTGTAGCCCTTGTCCAGGCGCCCTTCAAAGGCGCTTGGAGACCTCT 1771
 QY 675 TTGAAGCTGGGGCATCCGCGCTGGCGCTGTGGGCCATCGTGTGCTCTCCGCTGTGCA 734
 Db 1772 TTGAGAGCTGGGGCATCCGCGCTTGTGTGTGGGCCATCGTGTACTCTCCGTA 1831
 QY 735 ATGGAAGTGTGTGATGACGCTGTGCTGGGCGGCGCTGCGCCCTGCGCGGCTCAAGT 794
 Db 1832 ACGGCTGTGTGTGATGAGTCTTGGCAGCGGACCCAGCGCGCTGTCCCGCGCTCAAGC 1891
 QY 795 TTGTGATGAGTGTGATGACGCGCAACCTTGACTGGCATTTCCCTGTGCGCTCTAG 854
 Db 1892 TTGTGATGAGTGTGATGACGCGCAACCTTGTGCGGCTTCTGCGGCTTCTGCTGTG 1951
 QY 855 CTTCACTGTGCGCGCAGTGTGAGTGTGCTGTGAGTGTGAGGCGCGCTGGAGAGCG 914
 Db 1952 CTTGATGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTG 2011
 QY 915 GGCTAGGCTGTGCGGCGCATCTGCTTCTGCGAGTACTTGGGTGGAGGCACTGGGTGTC 974
 Db 2012 GTCTGGGCTGTGCGGCGCATCTGCTTCTGCGGCTTCTGCGGCTTCTGCGGCTTCTG 2071
 QY 975 TGCTACATCTGCGCGCAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTG 1034
 Db 2072 TGCTACATCTGCGCGCAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTG 2131
 QY 1035 AGTC-CCCTCCCTGGGCGAGGCTTGGAGCAGGCTTCTAGGCTGTGCGCACTGGCAGGG 1093
 Db 2132 AGGCGCGCTCGCTTGGCGAGCTTGGCGCAGCGCAGCTTGGATGCTTGGCGCTGGCGGG 2191
 QY 1094 CTGCGCGCGCATCTGCGCTTGGCTTCTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTG 1153
 Db 2192 CTGCGCGCAGCATCTGCGCTTGGCTTCTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTG 2251
 QY 1154 CCCTACGCGCGCATCTGAGGCTCAGCCAG-CAGCGCTGGGCTTCAAGTGTGCGCTGGTGTAT 1212
 Db 2252 CCCTACGCGCGCATCTGAGGCTCAGCCAG-CAGCGCTGGGCTTCAAGTGTGCGCTGGTGTAT 2311
 QY 1213 GATGAATCTCTTCTTCTGCTGTGCTGGCGGCTTCTTACATCAAACTGTACTGTGACCT 1272
 Db 2312 GATGAATCTCTTCTTCTGCTGTGCTGGCGGCTTCTTACATCAAACTGTACTGTGACCT 2371
 QY 1273 GCGCGGCGCGCATCTTGTAGGCGCTGTGGGACTGCGCCATGTGTAGGCACTGTGGCTGGCT 1332
 Db 2372 GCCACGGGCTGACTTTGAGGCGCTGTGGGACTGCGCCATGTGTAGGCACTGTGGCTGGCT 2431
 QY 1333 CATCTTGTGAGAGGCTTCTTACTGTGCTGCGGCTTCTTCAAGTGTGCGCTTCAAGTGTG 1392
 Db 2432 CATCTTGTGAGAGGCTTCTTACTGTGCTGCGGCTTCTTCAAGTGTGCGCTTCAAGTGTG 2491
 QY 1393 GGGCTCTTCTGCTGACGCGCGAGGCGCTCAAGTGTGCTTCTGCTGTGCTGTGCTGTGCTGTG 1452
 Db 2492 GGGCTCTTCTGCTGACGCGCGAGGCTGTCAAGTGTGCTTCTGCTGTGCTGTGCTGTGCTGTG 2551
 QY 1453 GCTTGTGCTTCTTCAACCCATGTGTGCTGTGCTTCTTCAACCCATGTGTGCTGTGCTGTGCT 1512
 Db 2552 GCTTGTGCTTCTTCAACCCATGTGTGCTGTGCTTCTTCAACCCATGTGTGCTGTGCTGTGCT 2611

QY 1513 TCGCGGCTTCTGCGCGCGCGCAGCGGAGTCTCAGGCGCCCTAGCCTATGCTGCGGCGCGGGA 1572
 Db 2612 TCGCGGCTTCTGCGCGCGCGCCTCGGTCCCGAGGCGCCCTAGCCTACGCTGAGCGCGTGA 2671
 QY 1573 GCTGGAAGAAGTCTCTGTGATTTACCCAGGCGCTTGTAGCCTTCTGTGATGTGATCT 1632
 Db 2672 GCTGGAAGAAGTCTCTGTGATTTACCCAGGCGCTTGTAGCCTTCTGTGATGTGATCT 2731
 QY 1633 CATTTCTGGAAGTCTCTGAGCTGGGCGCGCGCCCTGGGCTGGAGACCTATGGCTTCCCGCTC 1692
 Db 2732 TATTTCTGGAAGTCTCTGAGCTGGGCGCGCGCCCTGGGCTGGAGACCTATGGCTTCCCGCTC 2791
 QY 1693 AGTCACCTCATCTCTGTCAAGCAGCGGCGCGCCCTGGGCTGGAGACCTATGGCTTCCCGCTC 1752
 Db 2792 AGTCACCTCATCTCTGTCAAGCAGCGGCGCGCCCTGGGCTGGAGACCTATGGCTTCCCGCTC 2851
 QY 1753 AGAGCCAGAGGGAACCACTTTTGGGAACCCCAACCCCTCCATGGATGGAGAACCTGTGCT 1812
 Db 2852 AGAGCTGATGGAACCAAGTTTGGGAACCCCAACCCCTCCATGGAGAACCTGTGCT 2911
 QY 1813 GAGGCGAGAGGATCTACGCGAGGCTGGAGGCTTGTCAAGGCGTGGCGGCTTTCAGCC 1872
 Db 2912 GAAGGCGAGGAGGACCACTTTTGGGAGGCTGTGGGCTTGTCCGTTGGGCTTTCAGCC 2971
 QY 1873 CTCTGGCTTGGCGCTTCTTCAACGCTGTAATATATCCCTCCCATTTCTTCTTCCCGCTC 1932
 Db 2972 CTCTGGCTTCTTCTTCCCTTCTTCAACGCTGTAATATATCCCTTCTTCTTCTTCCCGCTC 3031
 QY 1933 TCTTCCCTTCTTCTTCCCTTCTTCAACGCTGTAATATATCCCTTCTTCTTCTTCCCGCTC 1992
 Db 3032 CC-----AATGATGGCTGCTTATATAAAGAAAGCAACTC 3065
 QY 1993 AACTCAGCAGTGTGATCTATAGCAGGATGGCGCAGTACCTGGCTCCACTGATCACCTCT 2052
 Db 3066 CAATCTC-----ATAGCAAGTGGCGCAACCTCTGACTCCATTTGTT--CTCT 3111
 QY 2053 CTCTGTGACCATCACCAAGGCTTCTTGGCTGCTTCTTGGCTTCTTCTTGGCTTCTTCTCAGC 2112
 Db 3112 CTCAAGACCTTAAACATGAGTGTTCAGTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 3168
 QY 2113 TTCACTTGTATCTGGGCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 2172
 Db 3169 TTCACTTCT 3228
 QY 2173 CTTTGTCTCT 2232
 Db 3229 TTTGCTATGAGGAAAGGAGGAAAGCAAGAGAGTGAAGGTATGAGGCGCTTCAAGAGC 3288
 QY 2233 AGGCGCAGT-----GGACAGGAGACCTCACAGAGAAAGGCTTGGAGGCTGATT 2282
 Db 3289 CATGATCAGTAAGTGTGAGTGTGAGGAGTGTGAGGAGTGTGAGGAGTGTGAGGAGTGTGAGGAG 3348
 QY 2283 TCCGCTGTGATCATG-----ATAGGATACAAATATGTTTCCATGTA 2325
 Db 3349 TACCAAGAGATTTGGAGAGTCTCCCTGTGACATATAGATATATAATATGTTTGTGTT 3408
 QY 2326 CCATTAATCTTGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2385
 Db 3409 CCATTAATCTTGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3468
 QY 2386 ATTAAAAA 2393
 Db 3469 ATTGCATA 3476

RESULT 4
 BQ425502
 LOCUS
 DEFINITION
 AGENCOURT_7898175 NIH_MGC_67 Homo sapiens cdna clone IMAGE:6154083
 5', mRNA sequence.
 BQ425502
 VERSION
 BQ425502.1 GI:21120817
 EST.
 KEYWORDS

SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 771)
AUTHORS NIH-MGC <http://mgi.nci.nih.gov/>
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgabbs-remail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone Distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
Plate: LLAM13494 row: j column: 04
High quality sequence stop: 587.

FEATURES
source
Location/Qualifiers
1..771
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:6154083"
/tissue_type="retinoblastoma"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH MGC 67"
/note="Organ: eye; Vector: pCMV-SPORT6; Site 1: NotI;
Site 2: SalI; Cloned unidirectionally. Primer: Oligo dT.
Average insert size 1.75 Kb. Library constructed by Life
Technologies."

ORIGIN
Query Match 30.1%; Score 747.8; DB 5; Length 771;
Best Local Similarity 99.0%; Pred. No. 9.1e-172;
Matches 763; Conservative 0; Mismatches 7; Indels 1; Gaps 1;
Qy 1643 GCTTCTGAGCTGGCGGCCCCCTGGGCTGGAGACCTATGGCTTCCCTCAGTGACCCCTC 1702
Db 1 GCTTCTGAGCTGGCGGCCCCCTGGGCTGGAGACCTATGGCTTCCCTCAGTGACCCCTC 60
Qy 1703 ATCTCTGTACAGCCAGGGGCCCCAGGCTGGAGGGCAGCCATTTGTAGAGCCAGAG 1762
Db 61 ATCTCTGTACAGCCAGGGGCCCCAGGCTGGAGGGCAGCCATTTGTAGAGCCAGAG 120
Qy 1763 GGGAAACCACTTTGGGAACCCCAACCTCCATGGATGAGAACTGCTGCTGAGGGCAGAG 1822
Db 121 GGGAAACCACTTTGGGAACCCCAACCTCCATGGATGAGAACTGCTGCTGAGGGCAGAG 180
Qy 1823 GGATCTACGCCAGCAGGTGGAGGCTTGTACAGGGGTGGCGGCTTTTCAGCCCTCTGGCTTG 1882
Db 181 GGATCTACGCCAGCAGGTGGAGGCTTGTACAGGGGTGGCGGCTTTTCAGCCCTCTGGCTTG 240
Qy 1893 GCCTTTGCTTACAGGTGTAATATCCCTCCCAATTTCTTCTCCCTCTCTCTCTCTCTCT 1942
Db 241 GCCTTTGCTTACAGGTGTAATATCCCTCCCAATTTCTTCTCTCTCTCTCTCTCTCTCT 300
Qy 1943 CCTCTCTCCCTCGGTGAATGATGGCTTCTTAAACAAATACAAACAAACAACTCAGCA 2002
Db 301 CCTCTCTCCCTCGGTGAATGATGGCTTCTTAAACAAATACAAACAAACAACTCAGCA 360
Qy 2003 GTGTGATCTATAGCAGGATGGCCAGTACCTGGCTTCCACTGATCACTCTCTCTGTGAC 2062
Db 361 GTGTGATCTATAGCAGGATGGCCAGTACCTGGCTTCCACTGATCACTCTCTCTGTGAC 420
Qy 2063 CATCACCAACGGGTGCTTTGGGCTGCTTTCCCTTCCCTTCCCTTCCCTTCCCTTCCCTTGA 2122
Db 421 CATCACCAACGGGTGCTTTGGGCTGCTTTCCCTTCCCTTCCCTTCCCTTCCCTTCCCTTGA 480
Qy 2123 TACTGGGCTCTTCTCTGATGCTGAAGCTGTGGACAGAGCCTGGACTTTGTCTG 2182
Db 481 TACTGGGCTCTTCTCTGATGCTGAAGCTGTGGACAGAGCCTGGACTTTGTCTG 540

Qy 2183 CTTAAGGGAATGAGGGAAGTAAGACAGTGAAGGGGTGGAGGTTGATCAGGGCAGCAGT 2242
Db 541 CTTAAGGGAATGAGGGAAGTAAGACAGTGAAGGGGTGGAGGTTGATCAGGGCAGCAGT 600
Qy 2243 GGACAGGGAGACCTCAGACAGAGAAAGGCTGGAAGGTGATTTCCCGTGTGACTCATGGATA 2302
Db 601 GGACAGGGAGACCTCAGACAGAGAAAGGCTGGAAGGTGATTTCCCGTGTGACTCATGGATA 660
Qy 2303 GGATACAAATGTGTTCATGTACCATTAATCTTGACATATGCCATGATGATTAAGACTTCC 2362
Db 661 GGATACAAATGTGTTCATGTACCATTAATCTTGACATATGCCATGATGATTAAGACTTCC 720
Qy 2363 TATTAATAATAGC-TTGGAAAGAGATTAAAAAAGGCGGCCG 2412
Db 721 TATTAATAATAGCTTTGGAAAGAGATTAAAAAAGGCGGCCG 771

RESULT 5
AW970780/c
LOCUS AW970780 647 bp mRNA linear EST 01-JUN-2000
DEFINITION EST382863 MAGE resequences, MAGK Homo sapiens cDNA, mRNA sequence.
ACCESSION AW970780
VERSION AW970780.1 GI:8160625
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 647)
AUTHORS Hegde,P., Qi,R., Abernathy,K., Dharap,S., Gaspard,R., Gay,C.,
Holt,I.E., Saeed,A.I., Sharov,V., Lee,N.H., Yeatman,T.J. and
Quackenbush,J.
TITLE Assessment of gene expression patterns in a model of colon tumor
metastasis using a 19,200 element cDNA microarray
JOURNAL Unpublished (2000)
COMMENT Contact: John Quackenbush
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850, USA
Tel: 301 838 3528
Fax: 301 838 0208
Email: johnq@tigr.org
Plate: 282
Seq primer: Forward.
Location/Qualifiers
1..647
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone_lib="MAGE resequences, MAGK"
/note="Vector: pBluescriptSKm"

ORIGIN
Query Match 23.7%; Score 590.2; DB 2; Length 647;
Best Local Similarity 97.9%; Pred. No. 3.5e-133;
Matches 598; Conservative 0; Mismatches 13; Indels 0; Gaps 0;
Qy 1777 GAACCCCAACCCCTCCATGGATGGAGAACTGCTGCTGAGGGCAGAGGATCTAGCCAGC 1836
Db 611 GAGGCCCAACCCCTCCATGGATGGAGAAATTTGTTAAGGGCAGAGGATTTAGCCAGC 552
Qy 1837 AGGTGAGGCTTGTACAGGGGTGGCGGCTTTTCAGCCCTCTGCGCTTGGCTTTGCTTCA 1896
Db 551 AGGTGAGGCTTGTACAGGGGTGGCGGCTTTTCAGCCCTCTGCGCTTGGCTTTGCTTCA 492
Qy 1897 CGTGTAAATATCCCTCCCAATTTCTTCTTCCCTCTCTTCCCTTCTCTCTCTCTCT 1956
Db 491 AGTGTAAATATCCCTCCCAATTTCTTCTTCCCTCTCTTCCCTTCTCTCTCTCTCT 432
Qy 1957 GGTCGAATGATGGCTGCTTCTTAAACAAATACAAACAAACAACTCAGCAGTGTATCTATAGC 2016
Db 431 GGTGAATGATGGCTGCTTCTTAAACAAATACAAACAAACAACTCAGCAGTGTATCTATAGC 372

Qy	2017	AGGATGGCCCAAGTACCTGGGTCCACTGATCACCTCTCTCTCTGTGACCATACCAACGGGT	2076
Db	371	AGGATGGCCCAAGTACCTGGGTCCACTGATCACCTCTCTCTGTGACCATACCAACGGGT	312
Qy	2077	GCCTCTTGGGCTGGCTTTCCCTTGGGCTTCCCTCAGCTTCACTTGTATCACTTGGGCTCTTC	2136
Db	311	GCCTCTTGGGCTGGCTTTCCCTTGGGCTTCCCTCAGCTTCACTTGTATCACTTGGGCTCTTC	252
Qy	2137	CTTGTCTATGCTGAAGCTGTGGACGAGACCTGGACTTTTGTCTGCTTAAGGGGAATGA	2196
Db	251	CTTGTCTATGCTGAAGCTGTGGACGAGACCTGGACTTTTGTCTGCTTAAGGGGAATGA	192
Qy	2197	GGGAAGTAAAGACAGTGAAGGGGTGGAGGGTTGATCAGGGCACAGTGGACAGGGAGACCT	2256
Db	191	GGGAAGTAAAGACAGTGAAGGGGTGGAGGGTTGATCAGGGCACAGTGGACAGGGAGACCT	132
Qy	2257	CACAGAAAGGCGCTGGAGGTGATTTCCCGTGTGACTCATGGATAGGATACAAAATGTG	2316
Db	131	CACAGAAAGGCGCTGGAGGTGATTTCCCGTGTGACTCATGGATAGGATACAAAATGTG	72
Qy	2317	TTCCATGTACCAATTAATCTTTGACATATGCCATGCAATAAGACTTCCTATTAAAATAAGCT	2376
Db	71	TTCCATGTACCAATTAATCTTTGACATATGCCATGCAATAAGACTTCCTATTAAAATAAGCT	12
Qy	2377	TTGGAAGAGAT	2387
Db	11	TTGGAAGAGAT	1

RESULT 6	BF159363	982 bp	mRNA	linear	EST 30-OCT-2000
LOCUS	601770339f1	NCI_CGAP_Lu29	Mus musculus	cdna clone	IMAGE:398955'6',
DEFINITION	BF159363		mRNA sequence.		
ACCESSION	BF159363				
VERSION	BF159363.1	GI:11039462			
KEYWORDS	EST.				
SOURCE	Mus musculus (house mouse)				
ORGANISM	Mus musculus				
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
AUTHORS	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.				
TITLE	1 (bases 1 to 982)				
JOURNAL	NIH-MGC http://mgc.nci.nih.gov/ .				
COMMENT	National Institutes of Health, Mammalian Gene Collection (MGC)				
	Unpublished (1999)				
	Contact: Robert Strausberg, Ph.D.				
	Email: cgapsb@mail.nih.gov				
	Tissue Procurement: Gilbert Smith, Ph.D.				
	cdna Library Preparation: Life Technologies, Inc.				
	cdna Library Arranged by: The I.M.A.G.E. Consortium (LLNL)				
	DNA Sequencing by: Incyte Genomics, Inc.				
	Clone distribution: MGC clone distribution information can be				
	found through the I.M.A.G.E. Consortium/LLNL at:				
	http://image.llnl.gov				
	Plate: LLAM9199	row: n	column: 09		
	High quality sequence stop: 663.				

FEATURES
source

RESULT 7
BG863804
LOCUS
DEFINITION
602798701F1 NCI_CGAP_Mam4 Mus musculus cDNA clone IMAGE:4934280 5',
linear mRNA 819 bp
EST 29-MAY-2001
IMAGE:4934280 5',
mRNA sequence.
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
Mus musculus
Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 819)
NIH-MGC <http://mgc.nci.nih.gov/>.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
REFERENCE
AUTHORS
TITLE
JOURNAL

COMMENT Contact: Robert Strausberg, Ph.D.
Email: c9apbs-r@mail.nih.gov
Tissue Procurement: Lothar Hennighausen Ph.D., Priscilla Furth Ph.D.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
Plate: LLAM10863 row: a column: 01
High quality sequence stop: 758.
Location/Qualifiers
1. .599
/organism="Mus musculus"
/mol_type="mRNA"
/strain="NMRI"
/db_xref="taxon:10090"
/clone="IMAGE:4934280"
/tissue_type="tumor, gross tissue"
/dev_stages="5 months"
/lab_host="DH10B"
/clone_lib="NCI CGAP Mam4"
/note="Organ: mammary; Vector: pCMV-SPORT6; Site 1: Sall; Site 2: NotI; Cloned unidirectionally. Primer: Oligo dT. Library constructed by Life Technologies. Investigators providing samples: Lothar Hennighausen/Priscilla Furth, NIH Reference for transgenic model: Li et al., Cell Growth and Differentiation 7, 3-11 (1996)."

ORIGIN

Query Match 21.5%; Score 534.4; DB 4; Length 819;
Best Local Similarity 84.0%; Pred. No. 1.7e-119;
Matches 626; Conservative 0; Mismatches 117; Indels 2; Gaps 2;

QY 136 GGCCTCCACACACCGCATCTGGGAATGTGAGCTNGACACTTCAGCCAGCTGAGTCTC 195
Db 41 GGCCTCCACATACAGGATCAAGAAATTTGGTGC-AGATACCTTCAGCCAGCTGGGCTC 99

QY 196 CTTGCAAGCCCTGGATCTTAGCTGAAACGCCATCCGGTCCATCCACCTGAGCGCTTCTC 255
Db 100 CTTGCAAGCTTAGACCTGAGTTGGAATGCCATCCGTCGATCCACCTGAGGCTTTCTC 159

QY 256 CACCTGCACCTCCCTGGTCAAGCTGGACCTGACAGAACACACAGCTGACACACCTGCCCT 315
Db 160 AACCTTCGATCTTTGGTTAAGCTGGACCTGACTGACAAACAGCTGACACACCTGCCCT 219

QY 316 GGCTGGACTTGGGGCTTGATGCATCTGAAGCTCAAGGAAACCTTGCTCTCTCCAGGC 375
Db 220 GGCTGGGCTGGAGGCTGATGCACCTGAAAGCTCAAGGAACTTGGGCCCTGTCTCAGGC 279

QY 376 CTTCTCCAAGACAGTTTCCAAAACCTGAGGATCTCTGGAGTG-CCTTATGCTACCACT 434
Db 280 CTTCTCCAAGACAGTTTCCAAAACCTGAGGATCTCTGGAGTGCCCTACGCTACCACT 339

QY 435 GGTGTCTCTATGGATGTGTGCCAGCTTCTTCAAGCCCTCTGGGAGTGGGAGGCTGAAG 494
Db 340 GGTGTCTCTACGGCATCTGTGCCAGCTTCTTCAAGACCTCTGGGAGTGGGAGGCGGAG 399

QY 495 ACCTTCACCTTGATGATGAGGAGTCTTCAAAAAGCCCTCTGGGCTCTCTTGCACAGCAAG 554
Db 400 ACTTTTCATCAGAAAGAGGAGGACCAAGAGGCGCCCTGGGTCTCTCTTGTGGACAAAG 459

QY 555 CAGAAACCACTATACCAAGGACCTGGATGAGCTCAGCTGGAGATGGAGGACTCAAGC 614
Db 460 CTGAAACCACTATACCTAGACCTGGATGAGCTCCAGATGGGACAGAGGACTCAAGC 519

QY 615 CACACCCAGTGTCCAGTGTAGCCCTTACTCCAGGCCCCCTTCAAGCCCTGTGAGTACCTCT 674
Db 520 CAAACCCAGTGTCCAGTGTAGCCCTTGTTCAGGCCCCCTTCAAGCCCTGCCAGACCTCT 579

QY 675 TTGAAAGCTGGGGCATCCGCCCTGGCCGTGTGGGCCCATCGTGTGCTCTCCGTGCTCTGCA 734

Db 580 TTGAGAGCTGGGGCATCCGCCCTTGTGTGTGGGCATCGTGTACTCTCCGCTACTCTGTA 639

QY 735 ATGAGTCTGTCTGTGACCGTGTTCGCTGGCGGCTTGCCTCCCTCCCGGTCAGAGT 794

Db 640 ACGGCTGTGTCTGTGACAGTCTTTGCCAGGGACCCAGCGCTGTCCCGGTCAGAGC 699

QY 795 TTGTGTAGGTGGGATTCAGCGGCGCAACACCTTGACCTTGACCTTGCGCTTCTTAG 854

Db 700 TTGTGTGGGTGGGATTCAGCGGCGCAACCGCCTTACGGGCAATTCCTGTGGGTCTCTGGG 759

QY 855 CCTCAGTCGATGCCCTGACCTTTGG 879

Db 760 CCTCGTGACGCGCTTGACCTATGG 784

RESULT 8
AA292507 599 bp mRNA linear EST 16-MAY-1997
LOCUS zt30b10.r1 Soares ovary tumor NbHOT Homo sapiens cDNA clone
DEFINITION IMAGE:723835 5', mRNA sequence.
ACCESSION AA292507
VERSION AA292507.1 GI:1940494
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 599)
AUTHORS Hillier, L., Allen, M., Bowles, L., Dubuque, T., Geisler, G., Jost, S., Kucaba, T., Lacy, M., Le, N., Lennon, G., Marra, M., Martin, J., Moore, B., Schellenberg, K., Steptoe, M., Tan, F., Theising, B., White, Y., Wyllie, T., Waterston, R. and Wilson, R.
WashU-Merck EST Project 1997
Unpublished (1997)
Contact: Wilson RK
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu
This clone is available royalty-free through LLNL; contact the IMAGE Consortium (info@image.llnl.gov) for further information.
Seq primer: -28ml3 rev2 ET from Amersham
High quality sequence stop: 495.

FEATURES
source
1. 599
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:723835"
/sex="Female"
/tissue_type="ovarian tumor"
/lab_host="DH10B (ampicillin resistant)"
/clone_lib="Soares ovary tumor NbHOT"
/note="Organ: ovary; Vector: pT73D (Pharmacia) with a modified polylinker; Site 1: Not I; Site 2: Eco RI; 1st strand cDNA was primed with a Not I - oligo(dT) primer [5', TGTTACCAATCTGAGTGGGAGCGCGGCTTTTCTTTTCTTTT 3'], double-stranded cDNA was size selected, ligated to Eco RI adapters (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of a modified pT73 vector (Pharmacia). Library constructed by Bento Soares and M.Patina Bonaldo."

ORIGIN

Query Match 20.9%; Score 520.2; DB 1; Length 599;
Best Local Similarity 98.9%; Pred. No. 4.8e-116;
Matches 555; Conservative 0; Mismatches 3; Indels 3; Gaps 3;

QY 1832 CCACAGGTGAGGCTTGTTCAGGGGGTGGCGGCTTTCAGCCCTCTGGCTTGGCCTTCTCT 1891

Db 4 CCACAGGTGAGGCTTGTTCAGGGGGTGGCGGCTTTCAGCCCTCTGGCCTTGGCCTTCTCT 63


```
Qy 1922 CTCTTCCCTCTCTTCCCTTCTCTCTCCCTCGGTGAATGATGGCTCTCTTCTTAAAC 1981
Db 661 CCTCTCCCGTCC-----AATGATGGCTGTTTATAAAG 694
Qy 1982 AAATACACAAAACCTACAGAGTGATCTATAGCAGGATGCCAGTACCTGGCTCCAC 2041
Db 695 AAAGACAACTCCAACCTC-----ATAGCAAGATGGCCAAACACCTCTGACTCA 742
Qy 2042 TGATCAGCTCTCTCTGTGACATACCAAGGGTCTCTTGGCTGGCTTCCCTTGG 2101
Db 743 TTGTT--CTCTCTCCAGCCCTTAAACCAATGAGTGTCTTCAAGTCTTGTCTTGTGG 800
Qy 2102 CCTTCTCAGCTTCACTTGTATCTAGGCTCTTCTTGTCTATCTCTGAAGCTGTGGAC 2161
Db 801 CCT---TACGTTTCACTTTCACCTTGGGCTTCTGTGCAATCAATCTTCTGACAGA 857
Qy 2162 AGAGACCTGGACTTTTGTCTGCTTAAAGGAATGAGGGAAGTAAAGACAGTGAAGGGGTG 2221
Db 858 GGCCTGGAAATTTGCATAGGAGAAAGGAAAGCAACACAGTGAAGGTATATGGGC 917
Qy 2222 GAGGTTGATCAGGCAAGTG-----GACAGGAGACCTCAGAGAAAGGCT 2271
Db 918 CCTGACAGACCATGATCAGTAAGTGACAGAGTGTGGGGAGGTCTCAGACAGCATGAC 977
Qy 2272 GGAAGGTGATTTCCGTGTGACTCATGG-----ATAGGATCAAAATG 2314
Db 978 TGAAGACAACTACCAAGACATGGAGAGTCTCCCTGTGACATAGAATATAAATG 1037
Qy 2315 TGTTCATGTACCATTAATCTTGCATATGATGCATGCAATAAGACTTCTATTAATAAAG 2374
Db 1038 TGTTTGTGTTCCATTAATCTTGACCTATGCTGTGCCAAAGTGTCTCTGTTAAATACA 1097
Qy 2375 CTTTGGAGAGATTAAAA 2393
Db 1098 CTTTGGAGACATTGCATA 1116
```

```
RESULT 11
AW009250/c
LOCUS
DEFINITION AW009250 509 bp mRNA linear EST 10-SEP-1999
mRNA sequence.
ACCESSION AW009250
VERSION AW009250.1 GI:5858028
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 509)
NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
Unpublished (1997)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: Elias Campo, M.D., Michael R. Emmert-Buck,
M.D., Ph.D.
CDNA Library Preparation: M. Bento Soares, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
www-bio.llnl.gov/bbrp/image/image.html
Seq primer: -40UP from Gibco
High quality sequence stop: 451.
FEATURES
source
1..509
location=Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:2504224"
```

```
/sex="pooled"
/tissue_type="colon"
/lab_host="DH10B"
/clone_lib="NCI_CGAP_Co3"
/notes="vector: pT73D-Pac (Pharmacia) with a modified
polylinker; Site 1: Not I; Site 2: Eco RI; 1st strand cDNA
was prepared from 12 pooled bulk tumor samples and primed
with a Not I - oligo(dT) primer. Double-stranded cDNA was
ligated to Eco RI adaptors (Pharmacia), digested with Not
I and cloned into the Not I and Eco RI sites of the
modified pT73 vector. Library went through one round of
normalization."
```

ORIGIN

```
Query Match 20.4%; Score 507.4; DB 2; Length 509;
Best Local Similarity 99.8%; Pred. No. 6.3e-113;
Matches 508; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Qy 1888 TGCTTTACACGTGTAAATATCCCTCCCATTTCTTCTCTCCCTCTCTTCCCTTTCCTCT 1947
Db 509 TGCTTTACACGTGTAAATATCCCTCCCATTTCTTCTCTCCCTCTCTTCCCTTTCCTCT 450
Qy 1948 CTCCTCCCTCGGTGAATGATGCTGCTTTCTTAAACAAATACAAACCAAACTCAGCAGTGTG 2007
Db 449 CTCCTCCCTCGGTGAATGATGCTGCTTTCTTAAACAAATACAAACCAAACTCAGCAGTGTG 390
Qy 2008 ATCTATAGAGGATGGCCAGTACCTGGCTCCACTGATCCTCTCTCTGTGACCATCA 2067
Db 389 ATCTATAGAGGATGGCCAGTACCTGGCTCCACTGATCCTCTCTCTGTGACCATCA 330
Qy 2068 CCAACGGGTGCTCTTGGCTGCTTCCCTTGGCTTCTCTCAGCTTCACTTCTGCTTAA 2127
Db 329 CCAACGGGTGCTCTTGGCTGCTTCCCTTGGCTTCTCTCAGCTTCACTTCTGCTTAA 270
Qy 2128 GGCCTCTTCTTGTCTGATGCTGAAGCTGTGACGACAGACCTGAGCTTTTGTCTGCTTAA 2187
Db 269 GGCCTCTTCTTGTCTGATGCTGAAGCTGTGACGACAGACCTGAGCTTTTGTCTGCTTAA 210
Qy 2188 GGGAAATGAGGAATGAACAGACAGTGAAGGGTGGAGGGTGGATCAGGCGACAGTGACCA 2247
Db 209 GGGAAATGAGGAATGAACAGACAGTGAAGGGTGGAGGGTGGATCAGGCGACAGTGACCA 150
Qy 2248 GGGAGACCTCACAGAGAAAGCCCTGGAAGGTGATTTCCCTGCTGACTCATGGATAGGATA 2307
Db 149 GGGAGACCTCACAGAGAAAGCCCTGGAAGGTGATTTCCCTGCTGACTCATGGATAGGATA 90
Qy 2308 CAAAATGTGTTCCATGTACCAATTAATCTTGACATATGCCATGATGATGATGATGATGATG 2367
Db 89 CAAAATGTGTTCCATGTACCAATTAATCTTGACATATGCCATGATGATGATGATGATGATG 30
Qy 2368 AATAAGCTTTGGAAGAGATTAAAAAAA 2396
Db 29 AATAAGCTTTGGAAGAGATTAAAAAAA 1
```

RESULT 12

```
BB636845
LOCUS
DEFINITION BB636845 RIKEN full-length enriched, adult male aorta and vein Mus
musculus cDNA clone A530037C04 5', mRNA sequence.
ACCESSION BB636845
VERSION BB636845.1 GI:16472635
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 690)
Arakawa,T., Carninci,P., Fukuda,S., Furuno,M., Hanagaki,T.,
Hara,A., Hiramoto,K., Hori,F., Ishii,Y., Ito,M., Kawai,J.,
Konno,H., Kouda,M., Koya,S., Matsuyama,T., Miyazaki,A., Nomura,K.,
Ohno,M., Okazaki,Y., Okido,T., Saito,R., Sakai,C., Sakai,K.,
Sano,H., Sasaki,D., Shibata,K., Shinagawa,A., Shiraki,T.,
```


Email: bento-soares@uiowa.edu
Tissue Procurement: Dr. Gregg Hageman
cDNA Library Preparation: Dr. M. Bento Soares, University of Iowa
cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: Researchers may obtain clones from Research Genetics (www.resgen.com).
The following repetitive elements were found in this cDNA
sequence: 1-21, >AT_rich#low_complexity
Seq primer: M13 Forward
POLYA=Yes.

FEATURES
source
1. 482
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="UI-E-DX1-agx-c-12-0-UI"
/tissue_type="fetal eyes"
/dev_stages="fetal"
/lab_hosts="DH10B (Life Technologies) (T1 phage resistant)"
/clone_lib="UI-E-DX1"
/note="Organ: eye; Vector: p7T73-Pac (Pharmacia) with a modified polylinker; Site 1: EcoR I; Site 2: Not I; UI-E-DX1 is a normalized cDNA library containing the following tissue(s): fetal eyes. The library was constructed according to Bonaldo, Lennon and Soares, Genome Research, 6:791-806, 1996. First strand cDNA synthesis was primed with an oligo-dr primer containing a Not I site. Double stranded cDNA was ligated to an EcoR I adaptor, digested with Not I, and cloned directionally into p7T73-Pac vector. The oligonucleotide used to prime the synthesis of first-strand cDNA contains a library tag sequence that is located between the Not I site and the (dr)18 tail. The sequence tag for this library is AGAATCAAGA. This library was created for the program, Gene Discovery in the Visual System, supported by National Eye Institute (NEI).
TAG_TISSUE=human fetal eyes
TAG_LIB=UI-E-DX1
TAG_SEQ=AGAATCAAGA"

ORIGIN
Query Match 19.2%; Score 477.4; DB 4; Length 482;
Best Local Similarity 99.8%; Pred. No. 1.4e-105;
Matches 478; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Qy 1925 TTCCCTCTCTCCCTTTCTCTCTCCCTCGGTGAATGATGGTCTTCTAAACAAA 1984
Db 482 TTCCCTCTCTCCCTTTCTCTCTCCCTCGGTGAATGATGGTCTTCTAAACAAA 423
Qy 1985 TACACCAAACTCAGCAGTGTATCTATAGCAGATGGCCAGTACTGCTCCACTGA 2044
Db 422 TACACCAAACTCAGCAGTGTATCTATAGCAGATGGCCAGTACTGCTCCACTGA 363
Qy 2045 TCACCTCTCTCTGTGACCATCACACGGGTGCTCTTGCCCTGGCTTTCCCTTGGCT 2104
Db 362 TCACCTCTCTCTGTGACCATCACACGGGTGCTCTTGCCCTGGCTTTCCCTTGGCT 303
Qy 2105 TCCTCAGCTTCACCTTGATCTAGTGGCCCTCTTCTTGTCATGCTCTGAAGCTGTGACACGA 2164
Db 302 TCCTCAGCTTCACCTTGATCTAGTGGCCCTCTTCTTGTCATGCTCTGAAGCTGTGACACGA 243
Qy 2165 GACCTGACTTTTGTCTGCTTAAGGGAATCAGGGAAGTAAAGACAGTGAAGGGTGGAG 2224
Db 242 GACCTGACTTTTGTCTGCTTAAGGGAATCAGGGAAGTAAAGACAGTGAAGGGTGGAG 183
Qy 2225 GGTGTATCAGGGCACAGTGGACAGGACCTCACAGAGAAAGCCTGGAAGGTGATTTC 2284
Db 182 GGTGTATCAGGGCACAGTGGACAGGACCTCACAGAGAAAGCCTGGAAGGTGATTTC 123
Qy 2285 CCGTGTGACTCATGGATAGGATACAAAATGTGTTTCCATGTACCAATTAATCTTGACATATG 2344
Db 122 CCGTGTGACTCATGGATAGGATACAAAATGTGTTTCCATGTACCAATTAATCTTGACATATG 63

Qy 2345 CCATGCATAAGACTTCCTATTAAATAAGCTTTGGAAGAGATTAAAAA 2403
Db 62 CCATGCATAAGACTTCCTATTAAATAAGCTTTGGAAGAGATTAAAAA 4
RESULT 14
BM993280/c 479 bp mRNA linear EST 17-JUN-2002
LOCUS UI-H-DT0-avi-i-10-0-UI.s1 NCI_CGAP_DT0 Homo sapiens cDNA clone
DEFINITION IMAGE:5879889 3', mRNA sequence.
ACCESSION BM993280
VERSION BM993280.1 GI:19712669
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 479)
AUTHORS NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index
JOURNAL Unpublished (1997)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-x@mail.nih.gov
Tissue Procurement: Dr. Jose Mercuende
cDNA Library Preparation: Dr. M. Bento Soares, University of Iowa
cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
Clone Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: Clone Distribution Information can be found through the I.M.A.G.E. Consortium/INL at: http://image.llnl.gov
The following repetitive elements were found in this cDNA
sequence: 1-21, >AT_rich#low_complexity
Seq primer: M13 FORWARD
POLYA=Yes.

FEATURES
source
1. 479
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5879889"
/tissue_type="Metastatic Chondrosarcoma"
/dev_stages="Adult"
/lab_hosts="DH10B (Life Technologies)"
/clone_lib="NCI_CGAP_DT0"
/note="Organ: Lung; Vector: p7T73-Pac (Pharmacia) with a modified polylinker; Site 1: EcoR I; Site 2: Not I; NCI CGAP DT0 is a cDNA library containing the following tissue(s): Metastatic Chondrosarcoma in Lung. The library was constructed according to Bonaldo, Lennon and Soares, Genome Research, 6:791-806, 1996. First strand cDNA synthesis was primed with an oligo-dr primer containing a Not I site. Double stranded cDNA was ligated to an EcoR I adaptor, digested with Not I, and cloned directionally into p7T73-Pac vector. The oligonucleotide used to prime the synthesis of first-strand cDNA contains a library tag sequence that is located between the Not I site and the (dr)18 tail. The sequence tag for this library is AACTGTTCGG.
TAG_TISSUE=lung metastatic chondrosarcoma
TAG_LIB=UI-H-DT0
TAG_SEQ=AACTGTTCGG"

ORIGIN
Query Match 19.1%; Score 476; DB 5; Length 479;
Best Local Similarity 100.0%; Pred. No. 3.1e-105;
Matches 476; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1928 CCCTCTCTCCCTTTCTCTCTCCCTCGGTGAATGATGGTCTTCTAAACAAATAC 1987
Db 479 CCCTCTCTCCCTTTCTCTCTCCCTCGGTGAATGATGGTCTTCTAAACAAATAC 420
Qy 1988 AACCAAACTCAGCAGTGTGATCTATAGCAGGATGGCCAGTACTGCTGCCACTGATCA 2047

```
Db 419 AACCAAACTCAGCAGTGTGATCTATAGCAGGATGCCCCAGTACCTGGCTCCACTGATCA 360
Qy CCTCTCTCTGTGACCATCACCAACGGGTGCCTTTGGCCCTGGCTTTCCCTTGGCCCTTCC 2107
Db CCTCTCTCTGTGACCATCACCAACGGGTGCCTTTGGCCCTGGCTTTCCCTTGGCCCTTCC 300
Qy 2108 TCAGGTCACCTTGATCTAGGCTCTTCCCTTGTGATGTCTGAAGCTGTGACACGAGAC 2167
Db 239 TCAGGTCACCTTGATCTAGGCTCTTCCCTTGTGATGTCTGAAGCTGTGACACGAGAC 240
Qy 2168 CTGAGCTTTTGTCTCTTAAGGGAATGAGGGAATGAAGCAGTGAAGGGGTGAGGGT 2227
Db 239 CTGAGCTTTTGTCTCTTAAGGGAATGAGGGAATGAAGCAGTGAAGGGGTGAGGGT 180
Qy 2228 TGATCAGGCGCAGTGAGGAGACCTTCACAGAGAAAGGCTTGAAGGTGATTTCCCG 2287
Db 179 TGATCAGGCGCAGTGAGGAGACCTTCACAGAGAAAGGCTTGAAGGTGATTTCCCG 120
Qy 2288 TGTGACTCATGATAGGATACAAATGTGTTCCATGTACCAATTAATCTTGACATATGCCA 2347
Db 119 TGTGACTCATGATAGGATACAAATGTGTTCCATGTACCAATTAATCTTGACATATGCCA 60
Qy 2348 TGCATAGAGCTTCTCTATTAATAAGCTTTCGAGAGATTAAAAAAGGAAAAA 2403
Db 59 TGCATAGAGCTTCTCTATTAATAAGCTTTCGAGAGATTAAAAAAGGAAAAA 4

RESULT 15
AA524536 C
LOCUS n945903.s1 NCI_CGAP_Co3 Homo sapiens cDNA clone IMAGE:937780 3',
DEFINITION mRNA sequence.
ACCESSION AA524536
VERSION AA524536.1 GI:2265464
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 633)
AUTHORS NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
JOURNAL Unpublished (1997)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: ccapbs-r@mail.nih.gov
Tissue Procurement: Elias Campo, M.D., Michael R. Emmert-Buck,
M.D., Ph.D.
cDNA Library Preparation: M. Bento Soares, Ph.D.
cDNA Library Arraying: Greg Lennon, Ph.D.
CDNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
www-bio.lnl.gov/bbrp/image/image.html
Insert Length: 619 Std Error: 0.00
Seq primer: -40ml3 fwd. ET from Amersham
High quality sequence stop: 371.
Location/Qualifiers
1..633
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:937780"
/sex="pooled"
/tissue_type="colon"
/lab_host="DH10B"
/clone_lib="NCI_CGAP_Co3"
/note="Vector: p77T3D-Pac (Pharmacia) with a modified
polylinker; Site 1: Not I; Site 2: Eco RI; 1st strand cDNA
was prepared from 12 pooled bulk tumor samples and primed
with a Not I - oligo(dT) primer. Double-stranded cDNA was
ligated to Eco RI adaptors (Pharmacia), digested with Not
```

```
I and cloned into the Not I and Eco RI sites of the
modified p77T3 vector. Library went through one round of
normalization. "

Query Match 19.1%; Score 474.2; DB 1; Length 633;
Best Local Similarity 98.8%; Fred. No. 9e-105;
Matches 509; Conservative 0; Mismatches 3; Indels 3; Gaps 3;

ORIGIN
Qy 1881 TGGCCTTTGCTTACACGTCGTAATAATATCCCTCCCATTTCTTCTTCCCTCTCTTCCCT 1940
Db 514 TTGGCTTGTCTTACACGTCGTAATAATATCCCTCCCATTTCTTCTTCCCTCTCTTCCCT 456
Qy 1941 TTCTCTCTTCC-CCCTCGGTGAATGATGGCTGCTTTCTAAAACAAATAACAAC-AAAACTC 1998
Db 455 TTCTCTCTTCCACCTCGGTGAATGATGGCTGCTTTCTAAAACAAATAACAACCAAACTC 396
Qy 1999 AGCAGTGTGATCTATAGCAGGATGGCCAGTACCTGGCTCCACTGATCACCTCTCTCTG 2058
Db 395 AGCAGTGTGATCTATAGCAGGATGGCCAGTACCTGGCTCCACTGATCACCTCTCTCTG 336
Qy 2059 TGACCATCACCAACGGGTGCCTCTTGGCTGGCTTTCCCTTGGCTTCCCTCAGCTTTCACC 2118
Db 335 TGACCATCACCAACGGGTGCCTCTTGGCTGGCTTTCCCTTGGCTTCCCTCAGCTTTCACC 276
Qy 2119 TTGATCTGGGCTCTTCTCTTGTCTGTAAGCTGTGGACGAGACCTGGACTTTTG 2178
Db 275 TTGATCTGGGCTCTTCTCTTGTCTGTAAGCTGTGGACGAGACCTGGACTTTTG 216
Qy 2179 TCTCTTAAGGGAATGAGGGAAGTAAAGACAGTGAAGGGGTGAGGGTTGATCAGGGCA 2238
Db 215 TCTCTTAAGGGAATGAGGGAAGTAAAGACAGTGAAGGGGTGAGGGTTGATCAGGGCA 156
Qy 2239 CAGTGGACAGGAGACCTTCACAGAGAAAGGCTTGAAGGTGATTTCCCGTGTGACTCATG 2298
Db 155 CAGTGGACAGGAGACCTTCACAGAGAAAGGCTTGAAGGTGATTTCCCGTGTGACTCATG 96
Qy 2299 GATAGGATACAAAATGTGTTCCATGTACCAATTAATCTTGACATATGCCATGCATAAAGAC 2358
Db 95 GATAGGATACAAAATGTGTTCCATGTACCAATTAATCTTGACATATGCCATGCATAAAGAC 36
Qy 2359 TTCTATTAATAAAGCTTTGGAAGAGATTAAAAA 2393
Db 35 TTCTATTAATAAAGCTTTGGAAGAGATTAAAAA 1

Search completed: April 14, 2005, 18:37:07
Job time : 5372.18 secs
```

This Page Blank (usp10)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 14, 2005, 11:59:23 ; Search time 286.868 Seconds
(without alignments)
14179.987 Million cell updates/sec

Title: US-10-664-667-4
Perfect score: 2486
Sequence: 1 taatacagctactatagg.....agtgctacctaattcaatt 2486

Scoring table: IDENTITY NUC
Gapop 10'0 , Gapext 1.0

Searched: 1202784 seqs, 818138359 residues

Total number of hits satisfying chosen parameters: 2405568

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_NA.*
1: /cgn2_6/ptodata/1/ina/5A_COMB.seq.*
2: /cgn2_6/ptodata/1/ina/5B_COMB.seq.*
3: /cgn2_6/ptodata/1/ina/6A_COMB.seq.*
4: /cgn2_6/ptodata/1/ina/6B_COMB.seq.*
5: /cgn2_6/ptodata/1/ina/PTUS_COMB.seq.*
6: /cgn2_6/ptodata/1/ina/backfiles.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES:

Result No.	Score	Query Match	Length	ID	Description
1	2255.8	90.7	2988	4	US-09-799-451-723 Sequence 723, App
2	439.4	17.7	453	3	US-09-397-787-240 Sequence 240, App
3	355.6	14.3	2724	4	US-09-170-496D-263 Sequence 263, App
4	355.6	14.3	4570	4	US-09-976-594-201 Sequence 201, App
5	352.4	14.2	2724	4	US-09-170-496D-277 Sequence 277, App
6	194	7.8	2612	4	US-09-495-050A-214 Sequence 214, App
7	194	7.8	4203	2	US-08-866-757-1 Sequence 1, Appli
8	194	7.8	4203	3	US-09-153-593-1 Sequence 1, Appli
9	181.8	7.3	723	4	US-09-495-050A-220 Sequence 220, App
10	105	4.2	2703	4	US-09-482-273-75 Sequence 75, Appli
11	101.8	4.1	473	2	US-08-866-757-3 Sequence 3, Appli
12	101.8	4.1	473	3	US-09-153-593-3 Sequence 3, Appli
13	100	4.0	1581	1	US-08-383-756-1 Sequence 1, Appli
14	100	4.0	1581	2	US-08-460-898-1 Sequence 1, Appli
15	95.4	3.8	1255	3	US-09-118-442-31 Sequence 31, Appli
16	95.4	3.8	1255	3	US-09-677-064-31 Sequence 31, Appli
17	94.6	3.8	2709	4	US-09-482-273-101 Sequence 101, Appl
18	94.2	3.8	1665	4	US-09-591-095-23 Sequence 23, Appli
19	86.6	3.5	826	4	US-09-620-312D-584 Sequence 584, App
20	86.6	3.5	907	4	US-09-620-312D-585 Sequence 585, App
21	85.6	3.4	2311	4	US-09-489-847-123 Sequence 123, App
22	85.4	3.4	1984	4	US-09-904-615-46 Sequence 46, Appli
23	82	3.3	1200	4	US-09-620-312D-564 Sequence 564, Appli
24	82	3.3	2544	3	US-09-483-371-1 Sequence 1, Appli
25	82	3.3	2544	4	US-09-957-156-1 Sequence 1, Appli
26	81.8	3.3	696	3	US-09-227-357-69 Sequence 69, Appli
27	81.6	3.3	2638	3	US-09-042-785A-22 Sequence 22, Appli

c 28	81	3.3	501	3	US-09-439-313-451	Sequence 451, App
c 29	81	3.3	501	3	US-09-352-616A-451	Sequence 451, App
c 30	81	3.3	501	4	US-09-636-215-451	Sequence 451, App
c 31	81	3.3	501	4	US-09-685-166A-451	Sequence 451, App
c 32	81	3.3	501	4	US-09-679-426-451	Sequence 451, App
c 33	81	3.3	501	4	US-09-759-143-451	Sequence 451, App
c 34	81	3.3	501	4	US-09-651-236-451	Sequence 451, App
c 35	81	3.3	706	3	US-09-439-313-449	Sequence 449, App
c 36	81	3.3	706	3	US-09-352-616A-449	Sequence 449, App
c 37	81	3.3	706	4	US-09-636-215-449	Sequence 449, App
c 38	81	3.3	706	4	US-09-685-166A-449	Sequence 449, App
c 39	81	3.3	706	4	US-09-679-426-449	Sequence 449, App
c 40	81	3.3	706	4	US-09-759-143-449	Sequence 449, App
c 41	81	3.3	706	4	US-09-651-236-449	Sequence 449, App
c 42	80.8	3.3	1898	4	US-09-589-510-7	Sequence 7, Appli
c 43	80.4	3.2	1006	3	US-08-911-423-3	Sequence 3, Appli
c 44	80.4	3.2	3327	3	US-08-689-421-26	Sequence 26, Appli
c 45	80.4	3.2	3327	3	US-09-389-528-26	Sequence 26, Appli

ALIGNMENTS

RESULT 1

US-09-799-451-723
; Sequence 723, Application US/09799451
; Patent No. 6783969
; GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom
; APPLICANT: Zhou, Ping
; APPLICANT: Goodrich, Ryle
; APPLICANT: Asundi, Vinod
; APPLICANT: Ren, Feiyan
; APPLICANT: Zhang, Jie
; APPLICANT: Xue, Aidong J.
; APPLICANT: Zhao, Qing A.
; APPLICANT: Wang, Jian-Rui
; APPLICANT: Ma, Yunqing
; APPLICANT: Yamazaki, Victoria
; APPLICANT: Chen, Rui-hong
; APPLICANT: Wang, Dunrui
; APPLICANT: Wang, Yonghong
; APPLICANT: Wehrman, Tom
; APPLICANT: Ghosh, Reena
; APPLICANT: Dmanac, Radoje T.
; TITLE OF INVENTION: No. 6783969el Nucleic Acids and
; TITLE OF INVENTION: Polypeptides
; FILE REFERENCE: 803
; CURRENT APPLICATION NUMBER: US/09/799,451
; CURRENT FILING DATE: 2001-03-05
; NUMBER OF SEQ ID NOS: 948
; SOFTWARE: pt_FL_genes Version 2.0
; SEQ ID NO 723
; LENGTH: 2988
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(2484)
US-09-799-451-723

Query Match	90.7%	Score	2255.8	DB	4	Length	2988
Best Local Similarity	99.9%	Pred. No. 0					
Matches	2268	Conservative	0	Mismatches	2	Indels	1
Gaps	1						
Qy	133	AGCGGCTCCACACACACCGCATCTGGGAATTTGGAGTNGACACCTTCAGCCAGCTGAG	192				
Db	718	ATCGGCTCCACACACACCGCATCTGGGAATTTGGAGT-GACACCTTCAGCCAGCTGAG	776				
Qy	193	CTCCCTCGAGCCCTGGATCTTAGCTGAGACGCCATCCGGTCCATCCACCTCAGGCGCTT	252				
Db	777	CTCCCTCGAGCCCTGGATCTTAGCTGAGACGCCATCCGGTCCATCCACCTCAGGCGCTT	836				

Qy	253	CTCCACCTGCACCTCCCTGGTCAAGCTGGAAGCTGACAGACCAACAGCTGACCAACTGCC	312
Db	837	CTCCACCTGCACCTCCCTGGTCAAGCTGGAAGCTGACAGACCAACAGCTGACCAACTGCC	896
Qy	313	CCTGCTGGACCTTGGGGGCTTGATGATCTGAAGCTCAAGGAACTTGCTCTCTCCCA	372
Db	897	CCTGCTGGACCTTGGGGGCTTGATGATCTGAAGCTCAAGGAACTTGCTCTCTCCCA	956
Qy	373	GGCTTCTCCAGGACAGTTTCCAAAACCTGAGGATCTGAGGAGCTTATGCTACCA	432
Db	957	GGCTTCTCCAGGACAGTTTCCAAAACCTGAGGATCTGAGGAGCTTATGCTACCA	1016
Qy	433	GTGCTGCTCCCTATGGGATGTGCGCAGCTCTTCAAGGCTCTGGGAGTGGGAGCTGA	492
Db	1017	GTGCTGCTCCCTATGGGATGTGCGCAGCTCTTCAAGGCTCTGGGAGTGGGAGCTGA	1076
Qy	493	AGACCTTCACTTGAATGAGGAGTCTTCAAAAAGGCTCTGGGCTCTTGGCAGACA	552
Db	1077	AGACCTTCACTTGAATGAGGAGTCTTCAAAAAGGCTCTGGGCTCTTGGCAGACA	1136
Qy	553	AGCAGAGAACCACTATGACAGGACCTGGATGAGCTCCAGCTGGAGATGGAGGACTCAAA	612
Db	1137	AGCAGAGAACCACTATGACAGGACCTGGATGAGCTCCAGCTGGAGATGGAGGACTCAAA	1196
Qy	613	GCCACACCCAGTGTCCAGTGTAGCCCTACTCCAGGCCCCCTCAAGCCCTGTGATACCT	672
Db	1197	GCCACACCCAGTGTCCAGTGTAGCCCTACTCCAGGCCCCCTCAAGCCCTGTGATACCT	1256
Qy	673	CTTTGAAAGCTGGGGCATCCGCTGGCCGTGTGGCCACTCGTGTGCTCTCCGTGCTCTG	732
Db	1257	CTTTGAAAGCTGGGGCATCCGCTGGCCGTGTGGCCACTCGTGTGCTCTCCGTGCTCTG	1316
Qy	733	CAATGGACTGTGTGCTGA CCGTGTTCGCTGGGGGCTCTGCCCTCTGCCCCCGGTCAA	792
Db	1317	CAATGGACTGTGTGCTGA CCGTGTTCGCTGGGGGCTCTGCCCTCTGCCCCCGGTCAA	1376
Qy	793	TTTTGTGTAGTGGATGCGATGCGAGCGCCAA CACCTTGACTGGCATTTTCTGTGGCTCT	852
Db	1377	TTTTGTGTAGTGGATGCGATGCGAGCGCCAA CACCTTGACTGGCATTTTCTGTGGCTCT	1436
Qy	853	AGCCTCAGTCAGTCCCTGACCTTTGGTCACTTCTGAGTACGAGAGCCCTGGGAGAC	912
Db	1437	AGCCTCAGTCAGTCCCTGACCTTTGGTCACTTCTGAGTACGAGAGCCCTGGGAGAC	1496
Qy	913	GGGGTAGGCTGCGGGCCA CTGGCTTCTTGGCAGTACTTGGGTGAGGAGCATCGGTGCT	972
Db	1497	GGGGTAGGCTGCGGGCCA CTGGCTTCTTGGCAGTACTTGGGTGAGGAGCATCGGTGCT	1556
Qy	973	GCTGCTACTTGGCGGAGTGAGTGAGAGCTCTCGGTCTCTGTGTCCGGGCTATGG	1032
Db	1557	GCTGCTACTTGGCGGAGTGAGTGAGAGCTCTCGGTCTCTGTGTCCGGGCTATGG	1616
Qy	1033	GAAGTCCCTTCTGGGACGCTTTCAGAGAGGGTCTTAGGCTGCTGGCAGTGGCAGG	1092
Db	1617	GAAGTCCCTTCTGGGACGCTTTCAGAGAGGGTCTTAGGCTGCTGGCAGTGGCAGG	1676
Qy	1093	GCTGGCGCGCACTGGCCCTGGGCTCAGTGGGAGAAATACGGGGCTCTCCCACTCTGCT	1152
Db	1677	GCTGGCGCGCACTGGCCCTGGGCTCAGTGGGAGAAATACGGGGCTCTCCCACTCTGCT	1736
Qy	1153	GCCCTAGCGGCACTGAGGGTCAAGCAGAGCCCTGGGCTTCAAGCTGGGCTTGGTAT	1212
Db	1737	GCCCTAGCGGCACTGAGGGTCAAGCAGAGCCCTGGGCTTCAAGCTGGGCTTGGTAT	1796
Qy	1213	GATGAACTCTTCTGTTTCTGGTCTGGCGGCTGCTTACATCAAACTGTACTGTGACCT	1272
Db	1797	GATGAACTCTTCTGTTTCTGGTCTGGCGGCTGCTTACATCAAACTGTACTGTGACCT	1856
Qy	1273	GCCGCGGGGCACTTTGAGGCGGTGTGGGACTGCGCCATGCTGAGGCACTGGGCTTGCT	1332
Db	1857	GCCGCGGGGCACTTTGAGGCGGTGTGGGACTGCGCCATGCTGAGGCACTGGGCTTGCT	1916

Qy	1333	CATCTTCGACAGCGGCTCTCTACTGTCCGTGGCTTCTCCTCAGCTTCGGCTCCATGCT	1392
Db	1917	CATCTTCGACAGCGGCTCTCTACTGTCCGTGGCTTCTCCTCAGCTTCGGCTCCATGCT	1976
Qy	1393	GGGCTCTTCCCTGTCAAGGCTGAGGCGCTCAAGTCTGTCTCCTGTGTGTGTGTGTGTGT	1452
Db	1977	GGGCTCTTCCCTGTCAAGGCTGAGGCGCTCAAGTCTGTCTCCTGTGTGTGTGTGTGTGT	2036
Qy	1453	GCCTGCTGCTCAAGGCTGAGGCGCTCAAGTCTGTCTCCTGTGTGTGTGTGTGTGTGTGT	1512
Db	2037	GCCTGCTGCTCAAGGCTGAGGCGCTCAAGTCTGTCTCCTGTGTGTGTGTGTGTGTGTGT	2096
Qy	1513	TCGGCGCTTCGGCGCCGCGCAGGAGCTCAGGCGCCCTAGCCTATGCTGGCGCGGGGA	1572
Db	2097	TCGGCGCTTCGGCGCCGCGCAGGAGCTCAGGCGCCCTAGCCTATGCTGGCGCGGGGA	2156
Qy	1573	GCTGGAGAAGAGCTCTGTGATTTACCCAGGCGCTGTGAGCTTCTCTGATGTGGATCT	1632
Db	2157	GCTGGAGAAGAGCTCTGTGATTTACCCAGGCGCTGTGAGCTTCTCTGATGTGGATCT	2216
Qy	1633	CATTCGGAAGCTTCTGAAAGCTGGCGCGCCCTCTGGGCTGGAGACCTATGGCTTCCCTC	1692
Db	2217	CATTCGGAAGCTTCTGAAAGCTGGCGCGCCCTCTGGGCTGGAGACCTATGGCTTCCCTC	2276
Qy	1693	AGTGACCTCATCTCTGTCTCAGCAGCAGGCGCGCCCTCAGGCTGGAGGCGAGCATTGTGT	1752
Db	2277	AGTGACCTCATCTCTGTCTCAGCAGCAGGCGCGCCCTCAGGCTGGAGGCGAGCATTGTGT	2336
Qy	1753	AGAGCCAGAGGGAACCACTTTGGGAAACCCCAACCCCTCCTATGATGGAGAACTGTGCT	1812
Db	2337	AGAGCCAGAGGGAACCACTTTGGGAAACCCCAACCCCTCCTATGATGGAGAACTGTGCT	2396
Qy	1813	GAGGCGAGAGGATCTACGCGCAGCAGGCTTGTCTCAGGCGGTGGCGCTTTCAGCC	1872
Db	2397	GAGGCGAGAGGATCTACGCGCAGCAGGCTTGTCTCAGGCGGTGGCGCTTTCAGCC	2456
Qy	1873	CTCTGGCTTGGCTTGTCTTACACGCTGAAATATCCTCCCACTTCTCTTCTTCCCTC	1932
Db	2457	CTCTGGCTTGGCTTGTCTTACACGCTGAAATATCCTCCCACTTCTCTTCTTCCCTC	2516
Qy	1933	TCCTTCCCTTCTCTCTCCCTCGGTGAAATGATGGCTTCTTAAACAAATAACAACA	1992
Db	2517	TCCTTCCCTTCTCTCTCCCTCGGTGAAATGATGGCTTCTTAAACAAATAACAACA	2576
Qy	1993	AACTCAGCAGTGTGATCTATAGCAGGATGGCGCAGTACCTGCTCAGCTGATCAGCTCT	2052
Db	2577	AACTCAGCAGTGTGATCTATAGCAGGATGGCGCAGTACCTGCTCAGCTGATCAGCTCT	2636
Qy	2053	CTCCTGTGACCATCACCAAGGCTCTTGGCTGGCTTCTCCCTTGGCTTCCCTCAGC	2112
Db	2637	CTCCTGTGACCATCACCAAGGCTCTTGGCTGGCTTCTCCCTTGGCTTCCCTCAGC	2696
Qy	2113	TTCACTTGTATCTGGGCTCTTCTTGTCTGATCTGTAAGCTGTGGAAGCTTGGACCTTGA	2172
Db	2697	TTCACTTGTATCTGGGCTCTTCTTGTCTGATCTGTAAGCTGTGGAAGCTTGGACCTTGA	2756
Qy	2173	CTTTTGTCTCTTAAAGGAAATGAGGAAATGAAAGCAGTGAAGGGGTGGAGGGTGTATC	2232
Db	2757	CTTTTGTCTCTTAAAGGAAATGAGGAAATGAAAGCAGTGAAGGGGTGGAGGGTGTATC	2816
Qy	2233	AGGCGACAGTGGACAGGAGACCTCAGAGAAAGGCTCGAAGGTGATTTCCCGTCTGA	2292
Db	2817	AGGCGACAGTGGACAGGAGACCTCAGAGAAAGGCTCGAAGGTGATTTCCCGTCTGA	2876
Qy	2293	CTCATGGATAGGATACAAAATGTGTCTCATGTACCTTAAATCTTGAATATGCGCATGCAT	2352
Db	2877	CTCATGGATAGGATACAAAATGTGTCTCATGTACCTTAAATCTTGAATATGCGCATGCAT	2936
Qy	2353	AAAGACTTCTTAAATAAAGCTTTGGAAAGAGATTAAAAAATAAAAAA 2403	
Db	2937	AAAGACTTCTTAAATAAAGCTTTGGAAAGAGATTAAAAAATAAAAAA 2987	

RESULT 2

US-09-397-787-240
; Sequence 240, Application US/09397787
; Patent No. 6468758
; GENERAL INFORMATION:
; APPLICANT: Benson, Darin R.
; APPLICANT: Lodes, Michael J.
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: King, Gordon E.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR OVARIAN
; TITLE OF INVENTION: CANCER THERAPY AND DIAGNOSIS
; FILE REFERENCE: 210121.466C2
; CURRENT APPLICATION NUMBER: US/09/397,787
; CURRENT FILING DATE: 1999-09-16
; NUMBER OF SEQ ID NOS: 334
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 240
; LENGTH: 453
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-397-787-240

Query Match 17.7%; Score 439.4; DB 3; Length 453;
Best Local Similarity 99.6%; Pred. No. 2.4e-99;
Matches 451; Conservative 0; Mismatches 1; Indels 1; Gaps 1;
Qy 1329 GGCTCATCTTCGACAGCGGGCTCCTACTGTCCCGTGGCCCTTCTCAGCTTCGCTCCA 1388
Db 1 GGCTCATCTTCGACAGCGGGCTCCTACTGTCCCGTGGCCCTTCTCAGCTTCGCTCCA 60
Qy 1389 TGCTGGGCTCTTCCCTGTCTACGCCCGAGCGCGTCAAGTCTGTCTGTGTGTGTGTGC 1448
Db 61 TGCTGGGCTCTTCCCTGTCTACGCCCGAGCGCGTCAAGTCTGTCTGTGTGTGTGTGC 120
Qy 1449 CCCTGGCTGCTGCTCAACCACTGTGTACCTGCTCTTCAACCCCACTTCGGGATG 1508
Db 121 CCCTGGCTGCTGCTCAACCACTGTGTACCTGCTCTTCAACCCCACTTCGGGATG 180
Qy 1509 ACCTTGGCGGCTTCGGCGCGCGGAGGAGTCAAGGCGGCTTAGCCCTATGCTGGCGCG 1568
Db 181 ACCTTGGCGGCTTCGGCGCGCGGAGGAGTCAAGGCGGCTTAGCCCTATGCTGGCGCG 240
Qy 1569 GGGAGCTGGAGAGAGCTCTGTGTATCTAACCAGGCGCTGGTACCTTCTGTATGTGG 1628
Db 241 GGGAGCTGGAGAGAGCTCCCGTGATCTAACCAGGCGCTGGTACCTTCTGTATGTGG 300
Qy 1629 ATCTCATCTCGAAGCTTCTGAAGCTGGCGCGCGGCGGCTGGGCTGGAGACCTATGCTTCC 1688
Db 301 ATCTCATCTCGAAGCTTCTGAAGCTGGCGCGCGGCGGCTGGGCTGGAGACCTATGCTTCC 360
Qy 1689 CCTCAGTGACCTCATCTCTGTGTAGCAGCGGCGGCGGCGGCTGGAGGCG-AGCCAT 1747
Db 361 CCTCAGTGACCTCATCTCTGTGTAGCAGCGGCGGCGGCGGCGGCTGGAGGCGAAGCCAT 420
Qy 1748 TGTGTAGGCGAGAGGGAACCACTTTGGGAAC 1780
Db 421 TGTGTAGGCGAGAGGGAACCACTTTGGGAAC 453

RESULT 3

US-09-170-496D-263
; Sequence 263, Application US/09170496D
; Patent No. 655339
; GENERAL INFORMATION:
; APPLICANT: Behan, Dominic P.
; APPLICANT: Chalmers, Derek T.
; APPLICANT: Liaw, Chen W.
; TITLE OF INVENTION: No. 655339-Endogenous, Constitutively Activated Human G Protein-
; TITLE OF INVENTION: Receptors
; FILE REFERENCE: AREN-0040
; CURRENT APPLICATION NUMBER: US/09/170,496D
; CURRENT FILING DATE: 1998-10-13
; NUMBER OF SEQ ID NOS: 294

; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 263
; LENGTH: 2724
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-170-496D-263

Query Match 14.3%; Score 355.6; DB 4; Length 2724;
Best Local Similarity 56.2%; Pred. No. 3.6e-78;
Matches 781; Conservative 0; Mismatches 580; Indels 28; Gaps 5;
Qy 136 GGCCTCCAAACACACACCGCATCTGGAAATTTGGAGCTTGACACCTTCAGCCAGCTGAGCTC 195
Db 1138 GACCTAAGACATAATGAATCTACGAAATTTAAAG-TTGACACTTTCAGCAGCTTGTCTTAG 1196
Qy 196 CCTCAAGCCCTGATCTTAGCTGGAACGCCATCCGGTCCATCCACCTCAGCCCTGAGGCTTCTC 255
Db 1197 CCTCCGATCGGTGAATTTGGCTTGGACAAATTTGCTATTATTACCCCAATGATTTTC 1256
Qy 256 CACCTCGACTCCCTGCTCAAGCTGAGCTGAGACCAAGCAACAGCTGAGCACTGCCCCCT 315
Db 1257 CACTTGGCATCCCTAATAAAGCTGGACCTATCGTCAACCTCCTGTCTGTCTTCTCTAT 1316
Qy 316 GGCTGGACTTGGGGCTTGATGATCTGAAGCTCAAAGGGAACCTTGTCTCTCTCCAGGC 375
Db 1317 AACTGGGTATCATGGTTAACTCACTTAAATTAACAGGAATCATGCTTACAGAGCTT 1376
Qy 376 CTCTCCAAAGACAGTTTCCCAAACTGAGGATCTCGAGGTGCTTATGCTTACCAGTG 435
Db 1377 GATATCATCTGAAACCTTCCAGAACTCAAGGTTATAGAAATGCTTATGTTACCAGTG 1436
Qy 436 CTGTCCCTATGGGATGTGTCCAGCTTCTTCAAGGCTCTGGGAGTGGGAGGCTGAAGA 495
Db 1437 CTGTGCATTTGGAGTGTGTGAGAAATGCTATAAGATTTCTAATCAATGAATAAAGGTGA 1496
Qy 496 CTTTACCTTGATGATGAGGAGTCTTCAAAAGGCCCTGGGCCCTCTTCCAGACAGC 555
Db 1497 CAACAGCAGTATGGACACCTTCAAGAAG- - - - -GATGCTGAAGTGT 1541
Qy 556 AGAGAACCACTATACACAGGACCTGGATGAGCTCAGCTGGAGAT- - -GGAGGACTCAA 612
Db 1542 TCAGGCTCAAGATGAACTGACCTTGAAGATTTCTGCTTGACTTTGAGGAAGACCTGAA 1601
Qy 613 GCCACACCCAGTGTCCAGTGTAGCCCTTACTTCCAGGCGGCTTCAAGCGCTCTGTAGTACCT 672
Db 1602 AGCCCTTCATTCAGTGTGAGTGTTCACCTTCCCGAGGCGGCTTCAAGCCCTGTGAACCT 1661
Qy 673 CTTTGAAGCTGGGGCATCCGCTGGCGGTGTGGCCATCGTGTGTCTCTCGTGTCTG 732
Db 1662 GCTTGATGGCTGGCTGATCAGAAATTTGGAGTGTGGACCATAGCAGTTCTGGCAGCTTACTTG 1721
Qy 733 CAATGGACTGTGTGCTGACCGGTTCGCTGGCGGCGCTGCCCGCTGCCCGGTCAG 792
Db 1722 TAATGCTTTGTGATCTCAACAGTTTTCAGATCC- - -CCTGTGATATTTCCCGCATTA 1778
Qy 793 GTTTGTGTAGTGTGATTTGAGCGGCGCAACACACTTGACTGGCATTTCTGTGGCTTCT 852
Db 1779 ACTGTTAATTTGGGTTCATCGCAGCAGTGAACATGCTCAGCGGAGTCTCCAGTGGCGTCT 1838
Qy 853 AGCCTCAGTGATGCCCTGACCTTTTGGTTCAGTTCTCTGAGTACGGAGCGGCGCTGGAGAC 912
Db 1839 GGCTGGTGTGATGCTTCACTTTTGGCAGCTTTTGCACAGACATGCTGCTGGTGGAGAA 1898
Qy 913 GGGCTAGGCTGGCGGCGCACTGGCTTCCCTGGCAGTACTTGGGTGGAGGATCGGTGCT 972
Db 1899 TGGGGTTGGTTGCCATGTCTATGTTTGTTCATTTTGTTCATTTTGTTCAGAAATCATCTGTTT 1958
Qy 973 GCTCTCAGCTGTGGCGGCGAGTGCAGGCTCTCCGCTCTCTGTGTCTGTGGGCTATGG 1032
Db 1959 CTTGCTTACTTGGCAGCGCTGGAGCGTGGGTCTCTGTGAAATATTTCTGCAAAATTTGA 2018
Qy 1033 GAAGTCCCGCTCCCTGGGCGAGCGTTTCGAGCAGGCGGTCTTAGCTGCGCTGGCAGCTGGCAGG 1092

QY 1333 CATCTTCGAGAGCGGCTCTCTACTGTGCTCCCGTGGCTTCTCCTCAGCTGCGCTCCATGCT 1392
Db 2574 GCTCTTCACCAACTCATGCTAAACTGCGCTGTGGCTTCTTGTCTCTCTCTCTCTTAA 2633
QY 1393 GGGCTCTTCCCTGTGACCGCGAGCGCTCAAGTCTGTCCTGCTGGTGGTGGTGGTGGT 1452
Db 2634 AAACCTTACATTTATCAGTCTCGAAGTAATTAAGTTTATCTCTCTGCTGGTGGTGGTGGT 2693
QY 1453 GCCTGCTGCTTCAACCCACCTGCTGTACCTGTCTTCAACCCACCTTCCGGGATGAGCT 1512
Db 2694 TCCTGCATGTCTCAATCCCTTCTCTACATCTTGTTCATCTCTCACTTAAAGGAGATCT 2753
QY 1513 TCGGCGGCT 1521
Db 2754 GGTGAGCT 2762

RESULT 5
US-09-170-496D-277
; Sequence 277, Application US/09170496D
; Patent No. 6555339
; GENERAL INFORMATION:
; APPLICANT: Behan, Dominic P.
; APPLICANT: Chalmers, Derek T.
; APPLICANT: Liaw, Chen W.
; TITLE OF INVENTION: No. 6555339-Endogenous, Constitutively Activated Human G Protein-
; FILE REFERENCE: AREN-0040
; CURRENT APPLICATION NUMBER: US/09/170,496D
; CURRENT FILING DATE: 1998-10-13
; NUMBER OF SEQ ID NOS: 294
; SOFTWARE: Patent In version 3.1
; SEQ ID NO 277
; LENGTH: 2724
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-170-496D-277

Query Match 14.2%; Score 352.4; DB 4; Length 2724;
Best Local Similarity 56.1%; Pred. No. 2.2e-77;
Matches 779; Conservative 0; Mismatches 582; Indels 28; Gaps 5;

QY 136 GGCCTCCACACACCGCATCTGGAAATTTGGAGCTNGACACCTTTCAGCCAGCTGAGCTC 195
Db 1138 GACCTTAAGACATAATGAAATCTACGAAATTAAG-TTGACACTTTCAGCAGCTGCTTAG 1196
QY 196 CTGCAAGCCCTGGATCTTAGCTGAAAGCCATCCGCTCCATCCACCTCAGGCTTCTC 255
Db 1197 CCTCGATCGGTGAATTTGGCTTGGAAACAAATTTGCTATTATTCACCCCAATGCATTTTC 1256
QY 256 CACCTGCACTCCCTGGTCAAGCTGAGCTGACAGACAAACAGCTGACACACTGCCCT 315
Db 1257 CACTTTGGCATCCCTTAATAAGCTGGACCTATCGTCCAACTCTCTGCTCTTTTCTTAT 1316
QY 316 GGCTGGACTTGGGGCTTGATGCACTCTGAAGCTCAAGGGAACCTTGTCTCTCCAGGC 375
Db 1317 AACTGGGTACATGTTTAACTCACTTAAATTAACAGAAATCATGCTTACAGAGCTT 1376
QY 376 CTCTCCAGAGACAGTTTCCCAAACTGAGATCTCGAGGTGCTTATGCTTACCACTG 435
Db 1377 GATATCATCTGAAACCTTCCAGAACTCAAGTTATAGAAATGCTTATGCTTACCACTG 1436
QY 436 CTGCTCCCTATGGATGTGTCAGCTTCTCAAGCCCTCTGGCAGTGGGAGGCTGAGA 495
Db 1437 GTGTGCAATTTGAGTGTGAGATGCTTAAGATTTCTAATCAATGAATAAAGGTGA 1496
QY 496 CTTTCACCTTGATGATGAGGATCTTCAAAAGGCGCTTCTTCCAGACCAAGC 555
Db 1497 CAACAGCAGTATGACGACCTTCAAGAA-----GATGCTGGATGTT 1541
QY 556 AGAGAACCACTATGACAGGACCTGGATGAGCTCCAGCTGGAGAT---GGAGGACTCAA 612
Db 1542 TCAGCTCAAGATGAACGTGACCTTGAAGATTTCTGCTTGAATTTGAGGAGGACCTGAA 1601

QY 613 GCCACACCCAGTGTCCAGTGTAGCCCTACTCCAGGCCCTTCAAGCCCTGTGAGTACCT 672
Db 1602 AGCCCTTCATTCAGTCAGTGTTCACCTTCCAGGCCCTTCAAACTCTGGAACACT 1661
QY 673 CTTTGAAGCTGGGGCATCCGCTGGCGTGTGGGCATCGTGTGCTCTCGTGTCTGTG 732
Db 1662 GCTTGTGCTGGCTGATCAGAAATGGAGTGTGACCATAGCAGTTCTGGCATTACTTGT 1721
QY 733 CAATGGACTGTGCTGACCGTGTTCGCTGGGGGCTGCCGCCCTGCCCGGTCAA 792
Db 1722 TAATGCTTGTGACTTCAACAGTTTTCAGATCC---CCTCTGTACATTTCCCCCATTA 1778
QY 793 GTTTGTGATAGTGTGAGTTGACGGCGCAACACCTTGACTGGCATTTCTGTGGCTTCT 852
Db 1779 ACTGTTAATTTGGGTTCATCGCAGCAGTGAACATGCTCACGGAGTCTCCAGTGGCTGT 1838
QY 853 AGCTCAGTCAGTCCCTGACCTTTGTGTCAGTTCTCTGAGTACGAGCCCTGGGAGAC 912
Db 1839 GGCTGGTGTGATGCTTCACTTTTGGCAGCTTTGACAGCATGCTGCTGTGGGAGAA 1898
QY 913 GGGCTAGGCTGCCGGCCACTGGCTTCTCGCAGTACTTGGTGGAGGATCGGTGCT 972
Db 1899 TGGGTGGTGGCATGTCTGTTTGTGCAATTTTGTCTCAGAAATCATCTGTTT 1958
QY 973 GCTGCTCACTTGGCCGAGTGCAGCTGTCGCTCTCGTCTCTCTGTGTCGGGCTATGG 1032
Db 1959 CTTGCTTACTTGGAGCCCTGGAGCTGGGTCTCTGTGAAATATTCGCAAAATTTGA 2018
QY 1033 GAATGCCCTTCTTGGGAGCGTTCGAGCAGGGTCTTAGGCTGCTTGGCAGTGGCAGG 1092
Db 2019 AACGAAAGCTCCATTTTCTAGCTGAAAGTAATCATTTTGTCTGTGCTGCTGCTG 2078
QY 1093 GCTGGCCGCGCACTGCCCTGGCTCAGTGGGAGAAATACGGGCGCTCCCACTCTGCT 1152
Db 2079 GACCATGGCCGAGTTCCTCTGCTGGTGGCAGCAAGTATGGCGCTCCCTCTCTGCT 2138
QY 1153 GCCCTAGCGCCACCTGAGGCTCAGCCAGCAGCCCTGGGCTTACCGTGGGCTGGTGTAT 1212
Db 2139 GCCTT-----TGCTTTTGGGAGCCAGCACCACCATGGCTACATGCTCTCATCTT 2192
QY 1213 GATGAATCTCTTCTGTTTCTGCTGGTGGCGGTGCTTACATCAAACTGTACTGTGACCT 1272
Db 2193 GCTCAATTCCTTGTCTCTCATGATGACCATTTGCTTACACCAAGCTCTACTGCAATTT 2252
QY 1273 GCCGGGGGCGACTTTGAGGCGGTGTGGAGTGGCCATGCTGAGGACAGTGGCTGCT 1332
Db 2253 GGAACAAGGAGACTGGAGAAATTTGGGACTGCTCTATGAAAAAACAACATTTGCCCTGT 2312
QY 1333 CATCTTCGACAGCGGCTCTCTACTGTCCGCTGGCTTCTCCTCAGCTTGGCTCCATGCT 1392
Db 2313 GCTCTTCAACCACTGCACTTAACTGCCCTGTGGCTTTCTTGTCTCTCTCTTAA 2372
QY 1393 GGGCTCTTCCCTGTACGCCCGAGCGCTCAAGTCTGTCTGCTGGTGGTGGTGGCT 1452
Db 2373 AAACCTTACATTTATCAGTCTGAAATAATTAAGTTTATCTCTGTGGTGGTGGCT 2432
QY 1453 GCCTGCTGCTCAACCCACTGCTGACCTCTTCAACCCCACTTCCGGGATGACCT 1512
Db 2433 TCCTGCAATGTCAATCCCTTCTCATCTTGTTCATCTCTCACTTTAAGAGGATCT 2492
QY 1513 TCGGCGGCT 1521
Db 2493 GGTGAGCT 2501

RESULT 6
US-09-495-050A-214
; Sequence 214, Application US/09495050A
; Patent No. 6492505
; GENERAL INFORMATION:
; APPLICANT: Roopa, Reddy
; APPLICANT: Guegler, Karl, J.

```

; APPLICANT: Au-Young, Janice
; TITLE OF INVENTION: COMPOSITION FOR DETECTION OF GENES ENCODING MEMBRANE-ASSOCIATED P
; FILE REFERENCE: PA-0013 US
; CURRENT APPLICATION NUMBER: US/09/495,050A
; CURRENT FILING DATE: 2000-01-31
; PRIOR APPLICATION NUMBER: 60/118,318
; PRIOR FILING DATE: February 1, 1999
; NUMBER OF SEQ ID NOS: 305
; SOFTWARE: PERL Program
; SEQ ID NO 214
; LENGTH: 2612
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No. 6492505 2470285CT1
US-09-495-050A-214

Query Match          7.8%; Score 194; DB 4; Length 2612;
Best Local Similarity 52.8%; Pred. No. 5,3e-38;
Matches 469; Conservative 0; Mismatches 410; Indels 9; Gaps 2;

Qy 623 AGTGTCCAGTGTAGCCCTTACTCCAGGCCCCCTTCAAGCCCTGTGAGTACCTCTTTGAAAGC 682
Db 1226 ATTATCCATTGTACACCTTCAACAGGTGCTTTTAAAGCCCTGTGAATATTTACTGGAAGC 1285

Qy 683 TGGGGCATCCCGCTGGCGGTGGGCCATCGTGTCTCCGTCTCGTCTGCAATGGACTG 742
Db 1286 TGSATGATTCGTCTTACTGTGTGTTCAATTTCTTGGTGTGCAATATTTTCAACCTGCTT 1345

Qy 743 GTGCTGTGACCGTGTTCGTCTGGCGGCGCTCCGCCCTCCCGGTCAGGTTTGTGGTA 802
Db 1346 GTTATTTTAAACAATTTGCATCTTG---TACATCAGTCCCTTCTGTCNAATTTTATA 1402

Qy 803 GGTGCGAFTGAGCGGCCAACACCTTGACTGGCAATTCCTGTGCGCTTCTAGCCTCAGTC 862
Db 1403 GGCTTGATTTCTGTGTCTAACTTATTCATGGGAATCTATAGTGCATCTTAACCTTTCTT 1462

Qy 863 GATGCCCTGACCTTTGTGTCAGTCTCTGAGTACGAGGCCCTGGGAGACGGGCTAGGC 922
Db 1463 GATGCTGTGCTCGGGCGAGATTCGCTGAATTTGGCAATTTGGTGGGAACCTGGCAGTGGC 1522

Qy 923 TGCCGGGCACTGGCTTCTCGCAGTACTTGGGTGGAGGCATCGGTGCTGCTCTCACT 982
Db 1523 TGCAAAAGTAGCTGGTCTTCTGAGTCTTCTCCTCAGAAAGTGCATATTTTATTAATG 1582

Qy 983 CTGCCCAGTGCAGTGCAGCGTCTCCGTCTCTCTGTGTCGGGGCTATGGGAAGTCCCCC 1042
Db 1583 CTAGCAACTGTGCAAAAGAGCTTATCTGCAAAAGATATATGAAATAATGGGAAGCAAT 1642

Qy 1043 TCCCTGGCAGCGTTCGAGCAGGGGTCTAGGCTGCCCTGGCACTGGCAGGGCTGGCGCC 1102
Db 1643 CATCTCAACAGTTTCCGGGTGTGCTGCCCCCTTTTGGCTTTCTTAGTGTCTACAGTACGAGC 1702

Qy 1103 GCATGCCCCCTGGCTCAGTGGGAGAAATACGGGCGCTCCGCCACTTGCCTCCCTACGCG 1162
Db 1703 TGTCTTCCCTTTTCCATAGAGGGGAATATCTGCATCACCCTTTGTTGGCAAT----- 1757

Qy 1163 CCACCTGAGGTACGACGAGCCCTGGGCTTACCGGTTCACCGTGGCCCTGGTGTGATGAATCC 1222
Db 1758 -TTCCTACAGGTGAAACGCCATCATATTAGGATTCATCTGTAAGCTTGTAGTCTATTAAACTCA 1816

Qy 1223 TCTGTCTTCTGTGTCGGCGGTGCTTACATCAAACTGACTGTGACTGTGACCTGCCCGGGGC 1282
Db 1816 -TTCCTACAGGTGAAACGCCATCATATTAGGATTCATCTGTAAGCTTGTAGTCTATTAAACTCA 1816

Qy 1817 CTAGCATTTTATTAATGGCGCTTATCTACACTAAGCTATATCTGCAACTTGGCAAAAGAG 1876
Db 1283 GACTTTGAGGCGGTGTGGGACTGCGCATGTGTGAGGACAGTGGCTGCTCATCTTCGCA 1342

Qy 1877 GACCTCTCAGNAAACTCACAATCTAGCATGATTAAGCATGTGCTTGGCTTAUCTTCACC 1936
Db 1343 GACGGGTCTCTACTGTCCGCTGGCCCTTCCCTAGCTTGGCCTCCACTGCTGGGCTCTTC 1402

Qy 1937 AATTGCATCTTTTCTGCCCCGTGGCGTTTTTTTTTCAITTTGACCATTGATCACTGCAATC 1996

```

```

Qy 1403 CCTGTACGCCCGAGGCGGTCAAGTCTGTCTCTGTGGTGGTGGTCCCTGCTGCTGCTGC 1462
Db 1997 TCTATCAGCCCCGAAATAATGAAGTCTGTACTCTGATATTTTTCATTTCCATTTGCTGCTGC 2056

Qy 1463 CTCAACCCACTGTGTACCTGCTCTTCAACCCCACTTCCGGGGATGAC 1510
Db 2057 CTGAATCCAGTCTGTATGTTTCTTCAACCCCAAGTTTAAAGAAGAC 2104

RESULT 7
US-08-866-757-1
; Sequence 1, Application US/08866757
; Patent No. 5858716
; GENERAL INFORMATION:
; APPLICANT: ELISHOURBAGY, NABIL A
; APPLICANT: LI, XIAOTONG
; APPLICANT: BERGSMAN, DEK J
; TITLE OF INVENTION: NOVEL 7TM RECEPTOR (H2CAA71)
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: RATNER & PRESTIA
; STREET: P.O. BOX 980
; CITY: VALLEY FORGE
; STATE: PA
; COUNTRY: USA
; ZIP: 19482
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/866,757
; FILING DATE: 30-MAY-1997
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: PRESTIA, PAUL F
; REGISTRATION NUMBER: 23,031
; REFERENCE/DOCKET NUMBER: GH-70055
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 610-407-0700
; TELEFAX: 610-407-0701
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4203 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cdna
US-08-866-757-1

Query Match          7.8%; Score 194; DB 2; Length 4203;
Best Local Similarity 52.8%; Pred. No. 6.4e-38;
Matches 469; Conservative 0; Mismatches 410; Indels 9; Gaps 2;

Qy 623 AGTGTCCAGTGTAGCCCTTACTCCAGGCCCCCTTCAAGCCCTGTGAGTACCTCTTTGAAAGC 682
Db 1117 ATTATCCATTGTACACCTTCAACAGGTGCTTTTAAAGCCCTGTGAATATTTACTGGGAGC 1176

Qy 683 TGGGGCATCCCGCTGGCGGTGGGCCATCGTGTGCTCTCCGTGCTCTGCAATGGACTG 742
Db 1177 TGGATGATTCGTCTTACTGTGTGTTTCAATTTCTTGGTGTGCAATATTTTCAACCTGCTT 1236

Qy 743 GTGCTGTGACCGGTGTGCTGGCGGCTGCCCCCTGCCCCGCTCAAGTTTGTGGTA 802
Db 1237 GTTATTTTAAACAATTTGCAATCTTG---TACATCACTGCTGCTGCTCAAAATTTTATA 1293

Qy 803 GGTGCGATTGACGCGGCCAACACCTTGACTGGGCAATTTCTGTGGCTTCTAGCCTCAGTC 862

```

Db 1294 GGCTGATTTCTGTGCTAACTTATTCATGGGAATCTATATGCGCATCTTAATTTTCTT 1353
Qy 863 GATGCCCTGACCTTTGGTTCAGTCTCTGAGTACGGAGCCGCTGGGACAGGGGTAGGC 922
Db 1354 GATGCTGTCTGGGGCAGATTGCTGATTTTGGCATTTGGTGGGAACCTGGCAGTGGC 1413
Qy 923 TGCCGGGCACCTGGCTTCTGGCAGTACTTGGGTGGAGGACATGGTCTGCTGCTACT 982
Db 1414 TGCAAACTGAACTGGGTTTCTGGCAGTTTCTCCTCAGAAAGTGCATATTTTATTATG 1473
Qy 983 CTGGCCGAGTGCAGTGCAGCTCTCCGTCTCCTGTGTCTGGGCTATGGGAAGTCCGCC 1042
Db 1474 CTAGCACTGCGAAAGAGCTTATCTGCAAAAGATATAATGAAAAATGGGAAGCAAT 1533
Qy 1043 TCCCTGGGCAGCTTTCAGAGCAGGGTCTAGGCTGCGCAGTGGCAGGGTGGCCGCC 1102
Db 1534 CATCTCAAAAGTTCGGGTGCTGCCCTTTGGCTTTCCTAGGTGCTACAGTAACAGC 1593
Qy 1103 GCATGCCCCCTGAGTGGGAGTAACGGGGCTCCCACTCTGCTGCCCTACGCG 1162
Db 1594 TGTTCCTCTTTCATAGAGGGGAATATCTGCATCACCCCTTTGTTGCCAT- ---- 1648
Qy 1163 CCACCTGAGGTACGACAGCCCTGGGCTTACCGTGGCCCTGGTGTGATGATGACTCC 1222
Db 1649 -TTCTCAGGTGAAACGCCATCATTAGSATTCATGTAACGTTAGTCTATTAACTCA 1707
Qy 1223 TTCTGTTCTCTGCTGCGCGGTCTACATCAACTGTACTGTGACCTGCGCGGGGC 1282
Db 1708 CTAGCATTTTATTAAGCGCGTATCTACACTAGCTATATCTGCAACTTGGAAAAGAG 1767
Qy 1283 GACTTTGAGCCGTGTGGGACTGCGCCATGTGGGACAGTGGCTGGCTCATCTTCGCA 1342
Db 1768 GACCTCTCAGAAAACCTCAACTAGCATGATTAAGCATGTCGCTTGGCTAATCTTCACC 1827
Qy 1343 GACGGCTCTCTACTGTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1402
Db 1828 AATTGCATCTTTTCTGCCCCCTGTCGCTTTTTCATTTGCACTTGTGCTGCTGCTGCT 1887
Qy 1403 CCTGTACGCCCGAGGCGCTCAAGTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1462
Db 1888 TCTATACGCCCGGAAATTAAGTCTGTTACTCTGATATTTTTCATGCTGCTGCTGCTGCT 1947
Qy 1463 CTCACCACTGCTGTACTCTCTTCAACCCCACTTCCGGGATGAC 1510
Db 1948 CTGAATCCAGTCTGTATGTTTCTTCAACCAAGTTTAAAGAAGAC 1995

RESULT 8

US-09-153-593-1
; Sequence 1, Application US/09153593A
; Patent No. 6174994
; GENERAL INFORMATION:
; APPLICANT: ELSHOURBAGY, NABIL A
; APPLICANT: LI, XIAOTONG
; APPLICANT: BERGSMAN, DEK J
; TITLE OF INVENTION: NOVEL 7TM RECEPTOR (H2CAA71)
; FILE REFERENCE: GH-70055-1
; CURRENT APPLICATION NUMBER: US/09/153.593A
; CURRENT FILING DATE: 1998-09-15
; EARLIER APPLICATION NUMBER: 08/866.757
; EARLIER FILING DATE: 1997-05-30
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1
; LENGTH: 4203
; TYPE: DNA
; ORGANISM: HOMO SAPIENS
US-09-153-593-1

Query Match 7.8%; Score 194; DB 3; Length 4203;
Best Local Similarity 52.8%; Pred. No. 6.4e-38;
Matches 469; Conservative 0; Mismatches 410; Indels 9; Gaps 2;

Qy 623 AGTGTCCAGTGTAGCCCTTACTCCAGGCCCTTCAAGCCCTGTGAGTACTCTTTGAAAGC 682
Db 1117 ATTATCCATTGTACACCTTCAACAGGTGCTTTTAAAGCCCTGTGAATATTTACTGGGAAGC 1176
Qy 683 TGGGGCATCCGCCCTGGCGGTGTGGCCATCGTGTGCTCTCCGTGCTCTGCAATGGAGCTG 742
Db 1177 TGSATGATTCGCTTACTGTGTGTTTCAATTTCTTGGTGTGATTAATTTTCAACCTGCTT 1236
Qy 743 GTGTGCTGACCGGTTCGTGGCGGGCTCCGCCCTGCCCCGCTCAAGTTTGTGTGTA 802
Db 1237 GTTATTTTAAACAATTTGCATCTTG- -TACATCACTGCTTCTGTCNAATTTGTTATA 1293
Qy 803 GGTGCGATTGAGCGGCCAACACCTTGACTGGCATTTCTGTGGCTTTCTAGCCCTCAGTC 862
Db 1294 GGCTTGATTTCTGTGCTAACTTATTCATGGGAATCTATCTGGCATCTTAATTTTCTT 1353
Qy 863 GATGCCCTGACCTTTGGTTCAGTCTCTGAGTACGGAGCCGCTGGGACAGGGGTAGGC 922
Db 1354 GATGCTGTGCTCGGGCAGATTGCTGAATTTGGCATTTGGGGAACCTGGCAGTGGC 1413
Qy 923 TGCCGGGCACCTGCTTCCCTGGCAGTACTTGGGTGGAGGCACTGGTCTGCTGCTCACT 982
Db 1414 TGCAAGTAACTGGTTCCTGAGTTCCTCTCAGAAAGTGCATATTTTATTATG 1473
Qy 983 CTGGCCGCGAGTGCAGCTCTCCGTCTCTCTGTGTCCGGGCTATGGGAAGTCCGCC 1042
Db 1474 CTAGCAACTGTCGAAAGAGCTTATCTGCAAAAGATATAATGAAAAATGGGAAGCAAT 1533
Qy 1043 TCCCTGGGCAGCTTTCAGAGAGGGGTCTAGGCTGCGCAGTGGCAGGGTGGCCGCC 1102
Db 1534 CATCTCAAAAGTTCGGGTGCTGCGCTTTTGGCTTTCCTAGGTGCTACAGTAACAGC 1593
Qy 1103 GCATGCCCCCTGCGCTCAGTGGGAGAAATACGGGCTCCGCCACTGCTGCCCTACGCG 1162
Db 1594 TGTTCCTCTTTCATAGAGGGGAATATCTGCATCACCCCTTTGTTGCCAT- ---- 1648
Qy 1163 CCACCTGAGGTACGACAGCCCTGGGCTTCAACCGTGGCCCTGGTGTGATGATGACTCC 1222
Db 1649 -TTCTCAGGTGAAACGCCATCATTAGGATTCATCTGTAAGCTATTAACTCA 1707
Qy 1223 TTCTGTTCTCTGCTGCGCGGTGCGCTTACATAAACTGTACTGTGACCTGCGGGGGC 1282
Db 1708 CTAGCATTTTATTAATGGCGTTTATCTACACTAAGCTATATCTGCAACTTGGAAAAGAG 1767
Qy 1283 GACTTTGAGCCGTGTGGGACTGCGCATGTGGGACAGTGGCTGCTCATCTTCGCA 1342
Db 1768 GACCTCTCAGAAAACCTCACAATCTAGCATGATTAAGCATGTCGCTTGGCTAATCTTCACC 1827
Qy 1343 GACGGCTCTCTACTGTCCTGCTGCGCTTCTCAGCTTCCGCTCCATGCTGGGCTCTTC 1402
Db 1828 AATTGCATCTTTTCTGCCCCCTGTCGCTTTTTCATTTGCACTTGTGCTGCTGCTGCT 1887
Qy 1403 CCTGTACGCCCGAGGCGCTCAAGTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1462
Db 1888 TCTATACGCCCGGAAATTAAGTCTGTTACTCTGATATTTTTCATGCTGCTGCTGCTGCT 1947
Qy 1463 CTCACCACTGCTGTACTCTCTTCAACCCCACTTCCGGGATGAC 1510
Db 1948 CTGAATCCAGTCTGTATGTTTCTTCAACCCCAAGTTTAAAGAAGAC 1995

RESULT 9

US-09-495-050A-220
; Sequence 220, Application US/09495050A
; Patent No. 6492505
; GENERAL INFORMATION:
; APPLICANT: Roopa, Reddy
; APPLICANT: Guegler, Karl, J.
; APPLICANT: Au-Young, Janice
; TITLE OF INVENTION: COMPOSITION FOR DETECTION OF GENES ENCODING MEMBRANE-ASSOCIATED F
; FILE REFERENCE: PA-0013 US
; CURRENT APPLICATION NUMBER: US/09/495.050A


```
; FILING DATE: 30-MAY-1997
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: PRESTIA, PAUL P
; REGISTRATION NUMBER: 23,031
; REFERENCE/DOCKET NUMBER: GH-70055
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 610-407-0700
; TELEFAX: 610-407-0701
; TELEX: 846169
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 473 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cdna
; US-08-866-757-3

Query Match          4.1%;      Score 101.8; DB 2; Length 473;
Best Local Similarity 56.8%; Pred. No. 2.1e-15;
Matches 187; Conservative 0; Mismatches 142; Indels 0; Gaps 0;

QY    1182 CAGCCCTGGGCTTCACCGTGGCCCTCGTGATGAAGAAGTCTCCTCTGTGTTCTCTGTGTCGTGG 1241
DB     ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY    1242 CCGGTGCCTACATCAAACTGACTGTGACCTGCCGCGGGGCGACTTTGAGGCGGTGTGGG 1301
DB     ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY    1302 ACTGCGGCATGGTGAGGACGTGGCTGCCTCATCTTCGACAGCGGCTCTCTCTACTGTC 1361
DB     ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY    134  AATCTAGCATGATTAAAGCATGTGCTTGGCTAATCTTCACAATGTCATCTTTTTCTGCC 193
DB     ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY    1362 CCGTGGGCTTCCCTCAGCTTCGCTCCATGCTGGGCTCTTCCCTGTCCAGCCCGAGGCGG 1421
DB     ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY    1422 TCAAGTCTGCTGCTGGTGGTGTGCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1481
DB     ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY    1482 TGCTCTTCACCCCCCACTTCGGGATGAC 1510
DB     ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY    314  TTCTTCAAACCCAAAAGTTTAAAGAGGAC 342
DB     ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

RESULT 12
US-09-153-593-3
; Sequence 3, Application US/09153593A
; Patent No. 6174994
; GENERAL INFORMATION:
; APPLICANT: ELISHOURBAGY, NABIL A
; APPLICANT: LI, XIAOTONG
; APPLICANT: BERGSMA, DERK J
; TITLE OF INVENTION: NOVEL 7TM RECEPTOR (H2CAA71)
; FILE REFERENCE: GH-70055-1
; CURRENT APPLICATION NUMBER: US/09/153,593A
; CURRENT FILING DATE: 1998-09-15
; EARLIER APPLICATION NUMBER: 08/866,757
; EARLIER FILING DATE: 1997-05-30
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 3
; LENGTH: 473
; TYPE: DNA
; ORGANISM: HOMO SAPIENS
; FEATURE:
; NAME/KEY: UNSURE
```

NAME: Donna E. Scherer
REGISTRATION NUMBER: 34,719
NAME: Carl J. Schwedler
REGISTRATION NUMBER: 36,924
REFERENCE/DOCKET NUMBER: CGNE 111
TELECOMMUNICATION INFORMATION:
TELEPHONE: (916) 753-6313
TELEFAX: (916) 753-1510
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 1581 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: CDNA to mRNA
US-08-383-756-1

Query Match 4.0%; Score 100; DB 1; Length 1581;
Best Local Similarity 100.0%; Pred. No. 1e-14;
Matches 100; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2387 TTAATAAAAAAAAAAGGCGCGCTCTAGAGGATCCAAGCTTACGTACGCGTGCATG 2446
Db 1477 TTAATAAAAAAAAAAGGCGCGCTCTAGAGGATCCAAGCTTACGTACGCGTGCATG 1536

Qy 2447 CGACGTACAGCTCTTCTATAGTGTACCTAAATTCATT 2486
Db 1537 CGACGTACAGCTCTTCTATAGTGTACCTAAATTCATT 1576

RESULT 14
US-08-460-898-1
Sequence 1, Application US/08460898
Patent No. 5850022
GENERAL INFORMATION:
APPLICANT: Dehesh, Katayoon
APPLICANT: Voelker, Toni Alois
APPLICANT: Hawkins, Deborah
APPLICANT: Davies, Huw Maelor
TITLE OF INVENTION: Production of Myristate in Plant Cells
NUMBER OF SEQUENCES: 17
CORRESPONDENCE ADDRESS:
ADDRESSEE: Calgene, Inc.
STREET: 1920 Fifth Street
CITY: Davis
STATE: CA
COUNTRY: USA
ZIP: 95616
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.50 inch, 2.0 MB
COMPUTER: Apple Macintosh
OPERATING SYSTEM: Macintosh 7.0
SOFTWARE: Microsoft Word 5.1(a)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/460,898
FILING DATE: 05-JUN-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/383,756
FILING DATE: 02-FEB-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/261,695
FILING DATE: 16-JUN-94
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/US93/10814
FILING DATE: 29-OCT-93
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: USSN 07/968,971
FILING DATE: 30-OCT-92
ATTORNEY/AGENT INFORMATION:

NAME: Elizabeth Lassen
REGISTRATION NUMBER: 31,845
NAME: Donna E. Scherer
REGISTRATION NUMBER: 34,719
NAME: Carl J. Schwedler
REGISTRATION NUMBER: 36,924
REFERENCE/DOCKET NUMBER: CGNE 111-1
TELECOMMUNICATION INFORMATION:
TELEPHONE: (916) 753-6313
TELEFAX: (916) 753-1510
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 1581 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: CDNA to mRNA
US-08-460-898-1

Query Match 4.0%; Score 100; DB 2; Length 1581;
Best Local Similarity 100.0%; Pred. No. 1e-14;
Matches 100; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2387 TTAATAAAAAAAAAAGGCGCGCTCTAGAGGATCCAAGCTTACGTACGCGTGCATG 2446
Db 1477 TTAATAAAAAAAAAAGGCGCGCTCTAGAGGATCCAAGCTTACGTACGCGTGCATG 1536

Qy 2447 CGACGTACAGCTCTTCTATAGTGTACCTAAATTCATT 2486
Db 1537 CGACGTACAGCTCTTCTATAGTGTACCTAAATTCATT 1576

RESULT 15
US-09-118-442-31
Sequence 31, Application US/09118442B
Patent No. 6197561
GENERAL INFORMATION:
APPLICANT: Martino-Catt, Susan J.
APPLICANT: Wang, Hongyu
APPLICANT: Beach, Larry R.
APPLICANT: Wang, Xun
APPLICANT: Bowen, Benjamin A.
TITLE OF INVENTION: Genes Controlling Phytate Metabolism in
TITLE OF INVENTION: Plants and Uses Thereof
FILE REFERENCE: 0706
CURRENT APPLICATION NUMBER: US/09/118,442B
CURRENT FILING DATE: 1998-07-17
EARLIER APPLICATION NUMBER: 60/055,446
EARLIER FILING DATE: 1997-08-11
EARLIER APPLICATION NUMBER: 60/055,526
EARLIER FILING DATE: 1997-08-08
EARLIER APPLICATION NUMBER: 60/053,944
EARLIER FILING DATE: 1997-07-28
NUMBER OF SEQ ID NOS: 31
SOFTWARE: Fast-Seq for Windows Version 3.0
SEQ ID NO 31
LENGTH: 1255
TYPE: DNA
ORGANISM: Zea mays
US-09-118-442-31

Query Match 3.8%; Score 95.4; DB 3; Length 1255;
Best Local Similarity 99.0%; Pred. No. 1.3e-13;
Matches 96; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 2387 TTAATAAAAAAAAAAGGCGCGCTCTAGAGGATCCAAGCTTACGTACGCGTGCATG 2446
Db 1159 TCAAAAAAAAAAAAAAAAAAGGCGCGCTCTAGAGGATCCAAGCTTACGTACGCGTGCATG 1218

Qy 2447 CGACGTACAGCTCTTCTATAGTGTACCTAAATTCATT 2483
Db 1219 CGACGTACAGCTCTTCTATAGTGTACCTAAATTCATT 1255

Search completed: April 14, 2005, 18:45:38
Job time : 290.868 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 14, 2005, 11:17:08 ; Search time 7111.6 Seconds
(without alignments)

16938.466 Million cell updates/sec

Title: US-10-664-667-4

Perfect score: 2486

Sequence: 1 taatcagctcactatagg.....agtgcacctaattcaatt 2486

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 4708233 seqs, 24227607955 residues

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl.*

1: gb_ba.*

2: gb_htg.*

3: gb_in.*

4: gb_on.*

5: gb_ov.*

6: gb_pat.*

7: gb_ph.*

8: gb_pl.*

9: gb_pr.*

10: gb_ro.*

11: gb_sts.*

12: gb_sy.*

13: gb_un.*

14: gb_vi.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2485	100.0	2486	6	AX301825 Sequence
2	2309.8	92.9	3451	6	CQ840814 Sequence
3	2309.4	92.9	2322	6	AX098217 Sequence
4	2257.4	90.8	2711	6	AX301828 Sequence
5	2247.8	90.4	3286	9	BC047905 Sequence
6	2247.8	90.4	3306	6	CQ840811 Sequence
7	2247.8	90.4	3349	6	AX926300 Sequence
8	2247.2	90.4	3325	9	AX358119 Sequence
9	2241	90.1	3429	6	AX451929 Sequence
10	2238	90.0	3492	6	AX301831 Sequence
11	2231.4	89.8	3273	6	CQ870634 Sequence
12	2219.6	89.3	3041	6	E42917 Novel gonad
13	2219.6	89.3	3041	6	AX016191 Sequence
14	2219.6	89.3	3381	6	AX527920 Sequence
15	2085.8	83.9	3119	6	AX549295 Sequence
16	2085.8	83.9	3119	9	AB049405 Homo sapi
17	1898	76.3	1899	6	AX301827 Sequence
18	1844.2	74.2	1987	6	E42916 Novel gonad
19	1844.2	74.2	1987	6	AX016189 Sequence

20	1754.4	70.6	2208	6	AX301830 Sequence
21	1754.4	70.6	2901	6	AX301833 Sequence
22	1752.6	70.5	2487	6	AX926297 Sequence
23	1743	70.1	119596	6	AX926299 Sequence
24	1743	70.1	177344	9	AL356953 Human DNA
25	1738.6	69.9	2484	9	AF190501 Homo sapi
26	1673	67.3	1681	6	BD159457 Primer fo
27	1673	67.3	1681	6	AX882015 Sequence
28	1673	67.3	1681	6	BD012484 Guanosine
29	1673	67.3	1681	9	AK027377 Homo sapi
30	1664.8	67.0	178757	2	AC074007 Homo sapi
31	1383.4	55.6	3637	6	AX301822 Sequence
32	1336	53.7	2901	6	AX301824 Sequence
33	1257.8	50.6	1272	9	AB083616 Homo sapi
34	1089.4	43.8	2044	10	BC026896 Mus muscu
35	987.2	39.7	253599	2	AC117827 Mus muscu
36	960.4	38.6	227401	2	AC105613 Rattus no
37	960.4	38.6	238582	2	AC098661 Rattus no
38	960.4	38.6	247017	2	AC110384 Rattus no
39	806.2	32.4	1827	6	CQ842816 Sequence
40	806.2	32.4	1827	9	AK123055 Homo sapi
41	723.2	29.1	1173	6	BD144308 Novel G-p
42	710.2	28.6	785	6	BD146757 Primer fo
43	710.2	28.6	785	6	AX866695 Sequence
44	513.4	20.7	515	6	E42913 Novel gonad
45	513.4	20.7	515	6	AX016183 Sequence

ALIGNMENTS

RESULT 1
AX301825
LOCUS AX301825 2486 bp DNA linear PAT 30-NOV-2001
DEFINITION Sequence 4 from Patent WO0185768.
ACCESSION AX301825
VERSION AX301825.1 GI:17382883
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS Gu, W.
TITLE Novel g-protein coupled receptors and uses therefor
JOURNAL Patent: WO 0185768-A 4 15-NOV-2001;
Millennium Pharmaceuticals, Inc. (US)
FEATURES
source
1..2486
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
CDS
22..1903
/note="unnamed protein product"
/codon_start=1
/protein_id="CAD13077.1"
/translation="NTTHYRESWACYRSGIPGSGTHASVERSOGLSIPAPASIALAL
AASNTASGKLELDTFQSLQSLQALDLSWNAIRSIHPEAFSTLSLVLKLDLTDNLQTT
LPALGLGLMHLKLNALQAFSKDSFKRILRLEVPYAYCCPYGMCASFFKASGG
WRADLHLDDEESKRLPLARQENHYDQDLDELQEMEDSKSPHSPVQSPPTGPGF
KPCBYLPESGIRLAWFAYVLLSVLCNGLVLTVPFAGGPAPLPVPKVFVGAAGANTL
TGTCGLASVDALTFCQFSGYRGWETGLGRATGFLAVLGSASVLLLTAAVQCS
VSVSCVRAYGKSPSLGSVRAGVLGCLALAGLALPLASVGEYCASPLCLPYAPEGQ
PALGLFTVALVMNNSFCFLVVGAYIKLYCDLPKRGDFEAWDCNWRHVAWLIFADGL
LYCFVAFSLFASMLGLFPVTPYAVKSVLLVLPACINPLLLYLPNHPHFDLRLRL
PRADSGFLAYAAAGLEKSCSDSTQNALVAFSDVDLILEASEAGRPQLETGYTFSVT
LISCOQPGAPRLGSHCVPEGNHFGNPQPSMDGELLRLAEGSTPAGGLSGGGGFPQ
SGLAFASHV"

ORIGIN

Query Match 100.0%; Score 2485; DB 6; Length 2486;
Best Local Similarity 100.0%; Pred. No. 0;

Matches 2486; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy	1	TAATACGACTC	1	TAATACGAGT	60
Db	1	TAATACGACTC	1	TAATACGAGT	60
Qy	61	GTCCAGCCAC	61	GTCCAGCCAC	120
Db	61	GTCCAGCCAC	61	GTCCAGCCAC	120
Qy	121	TCTTGTGCT	121	TCTTGTGCT	180
Db	121	TCTTGTGCT	121	TCTTGTGCT	180
Qy	181	CAGCAGTGA	181	CAGCAGTGA	240
Db	181	CAGCAGTGA	181	CAGCAGTGA	240
Qy	241	CCCTGAGGC	241	CCCTGAGGC	300
Db	241	CCCTGAGGC	241	CCCTGAGGC	300
Qy	301	GACCACTGC	301	GACCACTGC	360
Db	301	GACCACTGC	301	GACCACTGC	360
Qy	361	TGCTCTCC	361	TGCTCTCC	420
Db	361	TGCTCTCC	361	TGCTCTCC	420
Qy	421	TTATGCTTC	421	TTATGCTTC	480
Db	421	TTATGCTTC	421	TTATGCTTC	480
Qy	481	GTGGAGGCT	481	GTGGAGGCT	540
Db	481	GTGGAGGCT	481	GTGGAGGCT	540
Qy	541	CTTGTCCAG	541	CTTGTCCAG	600
Db	541	CTTGTCCAG	541	CTTGTCCAG	600
Qy	601	GGAGGACTC	601	GGAGGACTC	660
Db	601	GGAGGACTC	601	GGAGGACTC	660
Qy	661	CTGTGAGTA	661	CTGTGAGTA	720
Db	661	CTGTGAGTA	661	CTGTGAGTA	720
Qy	721	CTCGTGTCT	721	CTCGTGTCT	780
Db	721	CTCGTGTCT	721	CTCGTGTCT	780
Qy	781	GCCCCGGT	781	GCCCCGGT	840
Db	781	GCCCCGGT	781	GCCCCGGT	840
Qy	841	CTGTGGCTT	841	CTGTGGCTT	900
Db	841	CTGTGGCTT	841	CTGTGGCTT	900
Qy	901	CCGTGGGAG	901	CCGTGGGAG	960
Db	901	CCGTGGGAG	901	CCGTGGGAG	960
Qy	961	GGCATCGGT	961	GGCATCGGT	1020
Db	961	GGCATCGGT	961	GGCATCGGT	1020
Qy	1021	CCGGGCTAT	1021	CCGGGCTAT	1080
Db	1021	CCGGGCTAT	1021	CCGGGCTAT	1080

Qy	1081	GGCACTGGC	1081	GGCACTGGC	1140
Db	1081	GGCACTGGC	1081	GGCACTGGC	1140
Qy	1141	CCCACTCTG	1141	CCCACTCTG	1200
Db	1141	CCCACTCTG	1141	CCCACTCTG	1200
Qy	1201	GGCCCTGGT	1201	GGCCCTGGT	1260
Db	1201	GGCCCTGGT	1201	GGCCCTGGT	1260
Qy	1261	GTACTGTGA	1261	GTACTGTGA	1320
Db	1261	GTACTGTGA	1261	GTACTGTGA	1320
Qy	1321	CGTGGCTGG	1321	CGTGGCTGG	1380
Db	1321	CGTGGCTGG	1321	CGTGGCTGG	1380
Qy	1381	CGCCTCAT	1381	CGCCTCAT	1440
Db	1381	CGCCTCAT	1381	CGCCTCAT	1440
Qy	1441	GCTGCTGCC	1441	GCTGCTGCC	1500
Db	1441	GCTGCTGCC	1441	GCTGCTGCC	1500
Qy	1501	CCGGGATGA	1501	CCGGGATGA	1560
Db	1501	CCGGGATGA	1501	CCGGGATGA	1560
Qy	1561	TGGGCGGG	1561	TGGGCGGG	1620
Db	1561	TGGGCGGG	1561	TGGGCGGG	1620
Qy	1621	TGATGTGAT	1621	TGATGTGAT	1680
Db	1621	TGATGTGAT	1621	TGATGTGAT	1680
Qy	1681	TGCTTCCCT	1681	TGCTTCCCT	1740
Db	1681	TGCTTCCCT	1681	TGCTTCCCT	1740
Qy	1741	CAGCCATTG	1741	CAGCCATTG	1800
Db	1741	CAGCCATTG	1741	CAGCCATTG	1800
Qy	1801	AGAACTGCT	1801	AGAACTGCT	1860
Db	1801	AGAACTGCT	1801	AGAACTGCT	1860
Qy	1861	CGGCTTTCA	1861	CGGCTTTCA	1920
Db	1861	CGGCTTTCA	1861	CGGCTTTCA	1920
Qy	1921	TCTTCTCCC	1921	TCTTCTCCC	1980
Db	1921	TCTTCTCCC	1921	TCTTCTCCC	1980
Qy	1981	CAAAATACA	1981	CAAAATACA	2040
Db	1981	CAAAATACA	1981	CAAAATACA	2040
Qy	2041	CTGATCAC	2041	CTGATCAC	2100
Db	2041	CTGATCAC	2041	CTGATCAC	2100
Qy	2101	GCCTTCC	2101	GCCTTCC	2160
Db	2101	GCCTTCC	2101	GCCTTCC	2160

```
Qy 2161 CAGAGACCTGACCTTTTCTGCTTAAGGAAATGAGGAAAGTAAAGACAGTGAAGGGT 2220
Db 2161 CAGAGACCTGACCTTTTCTGCTTAAGGAAATGAGGAAAGTAAAGACAGTGAAGGGT 2220
Qy 2221 GGAGGGTTGATCAGGGACAGTGGACAGGAGACCTCAGAGAAAGGCGCTGGAGGTGA 2280
Db 2221 GGAGGGTTGATCAGGGACAGTGGACAGGAGACCTCAGAGAAAGGCGCTGGAGGTGA 2280
Qy 2281 TTTCCCGTGTGACTCATGGATAGATACAAAATGTGTTCCATGTACCAATTAATCTTGACA 2340
Db 2281 TTTCCCGTGTGACTCATGGATAGATACAAAATGTGTTCCATGTACCAATTAATCTTGACA 2340
Qy 2341 TATGCCATGCATAAAGACTTCTTATTAATAAGCTTTTGGAGAGATTAATAAAAAA 2400
Db 2341 TATGCCATGCATAAAGACTTCTTATTAATAAGCTTTTGGAGAGATTAATAAAAAA 2400
Qy 2401 AAAGGGCGCGCTCTAGAGGATCAAGCTTACGTACCGGTGCATGCGACCTCATAGCTC 2460
Db 2401 AAAGGGCGCGCTCTAGAGGATCAAGCTTACGTACCGGTGCATGCGACCTCATAGCTC 2460
Qy 2461 TTCTATAGTGTACCTAAATTAATT 2486
Db 2461 TTCTATAGTGTACCTAAATTAATT 2486

RESULT 2
CQ840814
LOCUS CQ840814 3451 bp DNA linear PAT 29-JUL-2004
DEFINITION Sequence 4 from Patent WO2004058818.
ACCESSION CQ840814
VERSION CQ840814.1 GI:50838404
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1
Muda, M.
Spliced variants of lgr6
Patent: WO 2004058818-A 4 15-JUL-2004;
Applied Research Systems ARS Holding N.V. (AN)
FEATURES
source
1. .3451
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
47. .1093
/note="unnamed protein product"
/codon_start=1
/protein_id="CAH05694.1"
/db_xref="GI:50838405"
/tranlation="MGRPRLTLVQVSIISARDLSMNNLTLOPLPHLRFLEELR
LSGNHLSHPQAFSGLYSLKLMNQNLGGIPAEALWELPSLSQSLRLDNLISLP
ERSPESSLSHMLDDNALTEIPVRLNLPALQMTALNRI SHIPDYAFQNLISL
VVLHNNRIQHLGTHSGPGLHLETLDLNKNLQEPFVAIRTLGRLOELGFHNNKTT
AIPKAFMNPGLQTHLYDNPIQFVGRSAFYLPKLHLISLNGAMDIQEPDLKGT
SLEILTFRAGIRLLPSGMCQQLPRLRVLELSHQIIBELPSLHRCQKLEMRSGGLYF
LPPSARGAWSGARV"

CDS
Query Match 92.9%; Score 2309.8; DB 6; Length 3451;
Best Local Similarity 99.7%; Pred. No. 0;
Matches 2325; Conservative 0; Mismatches 7; Indels 1; Gaps 1;

Qy 71 CGGTCCGTGGAGCGGAGCAGGGTCTGAGCCTGCCGCTCATCCAGCCTCTCTTGCTGCC 130
Db 1066 GGGGCGCATGGAGCGGAGCAGGGTCTGAGCCTGCCGCTCATCCAGCCTCTCTTGCTGCC 1125
Qy 131 CTAGCGGCTTCAACACAAACGCACTCTGGAAATTTGGAGCTNGACACCTTCAGCCAGCTG 190
Db 1126 CTAGCGGCTTCAACACAAACGCACTCTGGAAATTTGGAGCT-GACACCTTCAGCCAGCTG 1184
Qy 191 AGCTCCCTGCAAGCCCTGGATCTTAGCTGGAAACGCCATCCGGTCCATCCACCCTGAGGCC 250
```

ORIGIN

```
Db 1185 AGCTCCCTGCAAGCCCTGGATCTTAGCTGGAAACCCATCCGGTCCATCCACCCTGAGGCC 1244
Qy 251 TTCTCCACCTTGCACTCCCTGGTCAAGCTGACCTGACAGCAACACCAGCTGACCACTG 310
Db 1245 TTCTCCACCTTGCACTCCCTGGTCAAGCTGACCTGACAGCAACACCAGCTGACCACTG 1304
Qy 311 CCCCTGGCTGACATTGGGGGCTTGATGCATCTGAAGCTCAAAGGGAACCTTGCTCTCC 370
Db 1305 CCCCTGGCTGACATTGGGGGCTTGATGCATCTGAAGCTCAAAGGGAACCTTGCTCTCC 1364
Qy 371 CAGGCTTCTCAAGGACAGTTTCCAAAACCTGAGGATCCCTGGAGGTGCTTATGCTTAC 430
Db 1365 CAGGCTTCTCAAGGACAGTTTCCAAAACCTGAGGATCCCTGGAGGTGCTTATGCTTAC 1424
Qy 431 CAGTGTCTGCTTATGGATGTGCCAGCTTCTTCAAGGCTCTTGGGCACTGGGAGGCT 490
Db 1425 CAGTGTCTGCTTATGGATGTGCCAGCTTCTTCAAGGCTCTTGGGCACTGGGAGGCT 1484
Qy 491 GAAGACCTTCAACCTTGATGATGAGGAGTCTTCAAAAAGGCCCTTGGGCTCTCTTGCACA 550
Db 1485 GAAGACCTTCAACCTTGATGATGAGGAGTCTTCAAAAAGGCCCTTGGGCTCTCTTGCACA 1544
Qy 551 CAAGCAGAGAAACCACTATGACAGGACCTTGATGAGCTCCAGCTGGAGATGGAGGACTCA 610
Db 1545 CAAGCAGAGAAACCACTATGACAGGACCTTGATGAGCTCCAGCTGGAGATGGAGGACTCA 1604
Qy 611 AAGCCACACCCAGTGTCCAGTGTAGCCCTACTCCAGGCCCTTCAAGCCCTGTGAGTAC 670
Db 1605 AAGCCACACCCAGTGTCCAGTGTAGCCCTACTCCAGGCCCTTCAAGCCCTGTGAGTAC 1664
Qy 671 CTCTTTGAAAGCTGGGGCATCCGCTGCGGCTGTGGGCCATCTGTTGCTCTCCGTGCTC 730
Db 1665 CTCTTTGAAAGCTGGGGCATCCGCTGCGGCTGTGGGCCATCTGTTGCTCTCCGTGCTC 1724
Qy 731 TGCATGGAATGTGTGCTGTGACCGTGTTCGCTGGCGGCTCTGCCCTCTGCCCTCCGCTC 790
Db 1725 TGCATGGAATGTGTGCTGTGACCGTGTTCGCTGGCGGCTCTGCCCTCTGCCCTCCGCTC 1784
Qy 791 AAGTTTGTGTGTAGTGTGGATTGACAGGCCCAACACTTGTGCTGCACTTCTCTGTGGCCTT 850
Db 1785 AAGTTTGTGTGTAGTGTGGATTGACAGGCCCAACACTTGTGCTGCACTTCTCTGTGGCCTT 1844
Qy 851 CTAGCCTCAGTCGATGCCCTGACCTTTTGGTCACTTCTCTAGTAGCGGAGCCGCTGGGAG 910
Db 1845 CTAGCCTCAGTCGATGCCCTGACCTTTTGGTCACTTCTCTAGTAGCGGAGCCGCTGGGAG 1904
Qy 911 ACGGGCTAGGCTCCCGGGCCACTGGGCTTCCTGGCAGTACTTTGGGTGGAGGCACTCGGTG 970
Db 1905 ACGGGCTAGGCTCCCGGGCCACTGGGCTTCCTGGCAGTACTTTGGGTGGAGGCACTCGGTG 1964
Qy 971 CTGCTGTCTCACTCTGCGCGGAGTGCAGTGCAGCGCTCTCCGCTCTCTGTGTCGGGCTTAT 1030
Db 1965 CTGCTGTCTCACTCTGCGCGGAGTGCAGTGCAGCGCTCTCCGCTCTCTGTGTCGGGCTTAT 2024
Qy 1031 GGGAAAGTCCCCCTCCCTGGGCGAGGCTTCGAGCAGGGGTCTTAGGCTGCTCGGCACTGGCA 1090
Db 2025 GGGAAAGTCCCCCTCCCTGGGCGAGGCTTCGAGCAGGGGTCTTAGGCTGCTCGGCACTGGCA 2084
Qy 1091 GGGTGGCGCGCCGCACTGCCCCCTGGCTCAGTGGGAGAAATACGGGGCTCTCCCACTCTGC 1150
Db 2085 GGGTGGCGCGCCGCACTGCCCCCTGGCTCAGTGGGAGAAATACGGGGCTCTCCCACTCTGC 2144
Qy 1151 CTGCCCCACCGCCACCTGAGGGTTCAGCCAGAGCCTTGGGCTTTCACCGTGGGCTTGGTG 1210
Db 2145 CTGCCCCACCGCCACCTGAGGGTTCAGCCAGAGCCTTGGGCTTTCACCGTGGGCTTGGTG 2204
Qy 1211 ATGATGAACCTCTTCTGTTTCTGCTGTGCTGGCGGCTGTACATCAAACTGTACTGTGAC 1270
Db 2205 ATGATGAACCTCTTCTGTTTCTGCTGTGCTGGCGGCTGTACATCAAACTGTACTGTGAC 2264
Qy 1271 CTGCGCGGGCGGCACTTTGAGGCGCTGTGGGATGTGCGCCATGTGAGGCACTGGGCTGG 1330
```

[illegible]

RESULT 3	AX098217	Sequence 129 from Patent WO0118542.	2322 bp	DNA	linear	PAT 30-MAR-2001
LOCUS	AX098217					
DEFINITION	AX098217					
ACCESSION	AX098217					
VERSION	AX098217.1	GI:13515330				
KEYWORDS						
SOURCE						
ORGANISM	Homo sapiens (human)					
REFERENCE						
AUTHORS	Lee, J., Thompson, P. and Lillie, J.					
TITLE	Identification, assessment, prevention, and therapy of ovarian cancer					
JOURNAL	Patent: WO 0118542-A 129 15-MAR-2001;					
FEATURES	Millennium Predictive Medicine, Inc. (US)					
source	Location/Qualifiers					
	1..2322					
	/organism="Homo sapiens"					
	/mol_type="unassigned DNA"					
	/db_xref="taxon:9606"					
ORIGIN						
	Query Match	92.9%	Score 2309, 4;	DB 6;	Length 2322;	
	Best_Local Similarity	99.9%	Pred. No. 0;			
	Matches 2321; Conservative	0;	Mismatches	1;	Indels	1; Gaps
Qy	67	CCACGCTCCGTGGAGCGGAGCCAGGCTCTGAGCGCTGCGGCTCATCCAGCCCTCTCTTGC	126			
Db	1	CCACGCTCCGTGGAGCGGAGCCAGGCTCTGAGCGCTGCGGCTCATCCAGCCCTCTCTTGC	60			
Qy	127	TGCCCTAGCGGCTCCAAACACACCGCATCTGGGAAATTGGAGCTTGAGCTGACACCTTCAGCCA	186			
Db	61	TGCCCTAGCGGCTCCAAACACACCGCATCTGGGAAATTGGAGCTTGAGCTGACACCTTCAGCCA	119			
Qy	187	GCTGAGCTCCCTGCAAGCCCTGGATCTTAGCTGGAAAGCCCATCCGGTCCATCCAGCCCTGA	246			
Db	120	GCTGAGCTCCCTGCAAGCCCTGGATCTTAGCTGGAAAGCCCATCCGGTCCATCCAGCCCTGA	179			
Qy	247	GGCCTTCTCACCCCTGCACCTCCCTGCTGCAAGCTGGACCTGACAGACCAACAGCTGACCA	306			
Db	180	GGCCTTCTCACCCCTGCACCTCCCTGCTGCAAGCTGGACCTGACAGACCAACAGCTGACCA	239			
Qy	307	ACTGCCCTCGCTGGACTTCGGGCTTGATGTCATCTGAAGCTCAAAGGGAAACCTTGCTCT	366			
Db	240	ACTGCCCTCGCTGGACTTCGGGCTTGATGTCATCTGAAGCTCAAAGGGAAACCTTGCTCT	299			
Qy	367	CTCCAGGCTTCTTCCAAGCACAGTTTCCCAAACCTGAGGATCCTGAGGTGCTTTATGC	426			
Db	300	CTCCAGGCTTCTTCCAAGCACAGTTTCCCAAACCTGAGGATCCTGAGGTGCTTTATGC	359			
Qy	427	CTACCAGTGTGTCCTATGGGATGTGTGCAGCTTCTTCAAGCCCTCTGGGAGTGGGA	486			
Db	360	CTACCAGTGTGTCCTATGGGATGTGTGCAGCTTCTTCAAGCCCTCTGGGAGTGGGA	419			
Qy	487	GGCTGAAGACTTTCACCTTGATGATGAGGAGTCTTCAAAAAGGCCCTCTGGCCCTCCTTGC	546			
Db	420	GGCTGAAGACTTTCACCTTGATGATGAGGAGTCTTCAAAAAGGCCCTCTGGCCCTCCTTGC	479			
Qy	547	CAGACAAAGCAGAAACCACTATGACACGAGACCTGGATGAGCTCCAGCTGAGATGAGGA	606			
Db	480	CAGACAAAGCAGAAACCACTATGACACGAGACCTGGATGAGCTCCAGCTGAGATGAGGA	539			
Qy	607	CTCAAAGCCACACCCAGTGTCCAGTGTAGCCCTACTTCCAGGCCCCCTTCAAGCCCTGTGA	666			
Db	540	CTCAAAGCCACACCCAGTGTCCAGTGTAGCCCTACTTCCAGGCCCCCTTCAAGCCCTGTGA	599			
Qy	667	GTACCTCTTTGAAAGCTGGGCAATCCGCCCTGGCCGTGTGGCCCATCGTGTGCTCTCCGT	726			
Db	600	GTACCTCTTTGAAAGCTGGGCAATCCGCCCTGGCCGTGTGGCCCATCGTGTGCTCTCCGT	659			


```
QY 727 GCTCTGCAATGGAAGTGGTCTGCTGACCGTGTTCCTGCGGGCTGCCCCCTGCCCC 786
Db 660 GCTCTGCAATGGAAGTGGTCTGCTGACCGTGTTCCTGCGGGCTGCCCCCTGCCCC 719
QY 787 GGTCAAGTTTGTGGTAGGTGGATTTGCAGGCGCCAAACACCTTGACTGCGATTTCTGTGG 846
Db 720 GGTCAAGTTTGTGGTAGGTGGATTTGCAGGCGCCAAACACCTTGACTGCGATTTCTGTGG 779
QY 847 CTTCTAGCCTCAGTCGATGCCCCTGACCTTTGGTCAGTTCTCTGAGTAGCGAGCCCGCTG 906
Db 780 CTTCTAGCCTCAGTCGATGCCCCTGACCTTTGGTCAGTTCTCTGAGTAGCGAGCCCGCTG 839
QY 907 GGAGACGGGGTAGGCTGCGGGCCCACTGGCTTCTGCGAGTACTTGGGTGGAGGCATC 966
Db 840 GGAGACGGGGTAGGCTGCGGGCCCACTGGCTTCTGCGAGTACTTGGGTGGAGGCATC 899
QY 967 GGTGCTGCTGCTACTCTGGCGCAGTGCAGTGCAGCGTCTCCGTCCTCTGTGTCCGGC 1026
Db 900 GGTGCTGCTGCTACTCTGGCGCAGTGCAGTGCAGCGTCTCCGTCCTCTGTGTCCGGC 959
QY 1027 CTATGGGAAGTCCCCTCCCTGGGCGAGCGTTTCGAGCAGGGTCTTAGCTGCTGGCACT 1086
Db 960 CTATGGGAAGTCCCCTCCCTGGGCGAGCGTTTCGAGCAGGGTCTTAGCTGCTGGCACT 1019
QY 1087 GGCAGGGCTGCGCGCGCACTGCCCTTGGCCCTCAGTGGGAGAAATACGGGGCTCCCCACT 1146
Db 1020 GGCAGGGCTGCGCGCGCACTGCCCTTGGCCCTCAGTGGGAGAAATACGGGGCTCCCCACT 1079
QY 1147 CTGCTGCTGCTACGGCGCACCTGAGGGTTCAGCAGCAGCGCTTGGGCTTCACTGGGCCCT 1206
Db 1080 CTGCTGCTGCTACGGCGCACCTGAGGGTTCAGCAGCAGCGCTTGGGCTTCACTGGGCCCT 1139
QY 1207 GGTGATGATGAATCTCTTCTGTTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1266
Db 1140 GGTGATGATGAATCTCTTCTGTTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1199
QY 1267 TGACTGCTGCGGGGCGACTTTGAGGCGCTGTGGGACTGCGCCATGTTGAGGACAGCTGGC 1326
Db 1200 TGACTGCTGCGGGGCGACTTTGAGGCGCTGTGGGACTGCGCCATGTTGAGGACAGCTGGC 1259
QY 1327 CTGGCTCATCTTCGAGACGGGCTCTCTACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1386
Db 1260 CTGGCTCATCTTCGAGACGGGCTCTCTACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1319
QY 1387 CATGCTGGGCTCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1446
Db 1320 CATGCTGGGCTCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1379
QY 1447 GCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1506
Db 1380 GCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1439
QY 1507 TGACCTTCGGGGCTTCGGGCGCGGCGAGGGAAGTCTGAGGCGGCTTACGCTATGCTGGCG 1566
Db 1440 TGACCTTCGGGGCTTCGGGCGCGGCGAGGGAAGTCTGAGGCGGCTTACGCTATGCTGGCG 1499
QY 1567 CGGGAGCTGGAGAGAGCTCTGTTGATTCTACCCAGGCGCTGCTGCTGCTGCTGCTGCTG 1626
Db 1500 CGGGAGCTGGAGAGAGCTCTGTTGATTCTACCCAGGCGCTGCTGCTGCTGCTGCTGCTG 1559
QY 1627 GGATCTCATTTGGAAGCTTCTGGAAGCTGGGCGGCGGCTGCTGCTGCTGCTGCTGCTGCTG 1686
Db 1560 GGATCTCATTTGGAAGCTTCTGGAAGCTGGGCGGCGGCTGCTGCTGCTGCTGCTGCTGCTG 1619
QY 1687 CCCCTCAGTACCTCATCTCTGTCAGACGCGAGGCGGCGGCTGCTGCTGCTGCTGCTGCTG 1746
Db 1620 CCCCTCAGTACCTCATCTCTGTCAGACGCGAGGCGGCGGCTGCTGCTGCTGCTGCTGCTG 1679
QY 1747 TTGCTAGACGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1806
Db 1680 TTGCTAGACGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1739
```

```
QY 1807 GCTGCTAGGCGCAGAGGGATCTACGCCAGCAGGTGGAGGCTTGTACAGGGGTGGCGGCTT 1866
Db 1740 GCTGCTAGGCGCAGAGGGATCTACGCCAGCAGGTGGAGGCTTGTACAGGGGTGGCGGCTT 1799
QY 1867 TCAGCCCTCTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTT 1926
Db 1800 TCAGCCCTCTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTT 1859
QY 1927 CCCCTCTCTTCCCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1986
Db 1860 CCCCTCTCTTCCCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1919
QY 1987 CAACCAAACTCAGCAGTGTGATCTATAGCAGGATGGCCAGTACCTGGCTCCACTGATC 2046
Db 1920 CAACCAAACTCAGCAGTGTGATCTATAGCAGGATGGCCAGTACCTGGCTCCACTGATC 1979
QY 2047 ACCTCTCTCTCTGACCATCACCAACGAGTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTT 2106
Db 1980 ACCTCTCTCTCTGACCATCACCAACGAGTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTT 2039
QY 2107 CTCAGCTTTCACCTTTCATCTGGGCTCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2166
Db 2040 CTCAGCTTTCACCTTTCATCTGGGCTCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2099
QY 2167 CCTGAGCTTTTGTCTGCTTAAAGGAAATGAGGGAAGTAAAGACAGTGAAGGGGTGAGGG 2226
Db 2100 CCTGAGCTTTTGTCTGCTTAAAGGAAATGAGGGAAGTAAAGACAGTGAAGGGGTGAGGG 2159
QY 2227 TTGATCAGGCGCAGTGGACAGGAGGACCTCACAGAGAAAGGCTTGGAGGTGATTTCCC 2286
Db 2160 TTGATCAGGCGCAGTGGACAGGAGGACCTCACAGAGAAAGGCTTGGAGGTGATTTCCC 2219
QY 2287 GTGTGACTCATGGATAGGATACAAATGTGTTCATGTACCATTAATCTTGACATATGCC 2346
Db 2220 GTGTGACTCATGGATAGGATACAAATGTGTTCATGTACCATTAATCTTGACATATGCC 2279
QY 2347 ATGCATAAAGACTTCTCTATTAAATAAGCTTTGGAAGAGATTA 2389
Db 2280 ATGCATAAAGACTTCTCTATTAAATAAGCTTTGGAAGAGATTA 2322
```

```
RESULT 4
AX301828
LOCUS AX301828 2711 bp DNA linear PAT 30-NOV-2001
DEFINITION Sequence 7 from Patent WO0185768.
ACCESSION AX301828
VERSION AX301828.1 GI:17382887
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS Gu, W.
TITLE Novel g-protein coupled receptors and uses therefor
JOURNAL Patent: WO 0185768-A 7 15-NOV-2001;
Millennium Pharmaceuticals, Inc. (US)
FEATURES
source
1. .2711
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
<1..2211
/feature="unassigned protein product"
/codon_start=1
/protein_id="CAD13079.1"
/db_xref="GI:1738288"
/translation="GLNLETLDLNYKLOBPVARTLGRLOELGFHNNIKAIPEK
AFMNPILQITHEVDNPIQFVGRSAFOYLPKHLTSLNGAMDIQEPDLKCTTSLIL
TLFRGTRILPSCMCQQLPRLRVLELSHNQIEELPSLHRCQKLEIGLQHNRIWEIGA
TLFQSLQALDLSWNAIRSIHPEAFSLHSLVKLDLDTLTLPLAGLGGMLHLK
LKNLALQAFSKDSFPKRLILEVYAYCCPYGMCASFPAQSGWBAEDLHLDDES
SKRPLGLLARQAEHNDQLDLELQLEMEDSKPHPSVQCSPTTGPFPKFCYLFESWGR
```

```
FEATURES
source
1. .2711
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
<1..2211
/feature="unassigned protein product"
/codon_start=1
/protein_id="CAD13079.1"
/db_xref="GI:1738288"
/translation="GLNLETLDLNYKLOBPVARTLGRLOELGFHNNIKAIPEK
AFMNPILQITHEVDNPIQFVGRSAFOYLPKHLTSLNGAMDIQEPDLKCTTSLIL
TLFRGTRILPSCMCQQLPRLRVLELSHNQIEELPSLHRCQKLEIGLQHNRIWEIGA
TLFQSLQALDLSWNAIRSIHPEAFSLHSLVKLDLDTLTLPLAGLGGMLHLK
LKNLALQAFSKDSFPKRLILEVYAYCCPYGMCASFPAQSGWBAEDLHLDDES
SKRPLGLLARQAEHNDQLDLELQLEMEDSKPHPSVQCSPTTGPFPKFCYLFESWGR
```

CDS

LAVNAIIVLLSVCNGLVLLTVFAGGPAPLPVKKFVVGAIAGANTLITGSCGLLASVDA
LTFQFSEYGARWETGLCRATGFLAVLGSEASVLLLTAAVOCSSVSVCRAYGKSP
SLGSRVAGVLGCLALAGLALPLASVGEYGCASPLCLPYAPPEGOPALGFTVALVNM
NSPFLVVGAYIKLYCDLPRGDFEAVWDCMWRHVAWLIFADGLLYCPVAFSPASH
LGLFPVTPAEVKSIVLLVPLPUPACINPLLYLLFNPFRDDLRRLRPRAGSDGPIAYAA
AGELEKSCDSQVALVAFSPDVLJLEASERGPGLTFYFPSPVTLISCOOPGAPRLE
GSHCVPEGNHFNQPSMDGELLRLAEGSTPAGGSLSGGGGFPQPSGLAFASHV"

ORIGIN

Query Match 90.8%; Score 2257.4; DB 6; Length 2711;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 2269; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

Qy	133	AGCGGCTCCAAACACACCGCATCTGGGAATTCGAGCTNGACACCTTCAGCCAGCTGAG	192
Db	442	ATCGGCTCCAAACACACCGCATCTGGGAATTCGAGCT - GACACCTTCAGCCAGCTGAG	500
Qy	193	CTCGCTCAAGCCCTCGATCTTAGCTGGAACGCCATCCGGTCCATCCACCCCTGAGGCTT	252
Db	501	CTCGCTCAAGCCCTCGATCTTAGCTGGAACGCCATCCGGTCCATCCACCCCTGAGGCTT	560
Qy	253	CTCACCCCTGACCTCCCTGCTCAAGCTGGAGCTGACAGAACACAGCTGACCACTGCC	312
Db	561	CTCACCCCTGACCTCCCTGCTCAAGCTGGAGCTGACAGAACACAGCTGACCACTGCC	620
Qy	313	CCTGGCTGGACTTGGGGCTTGATGATCTCTGAAGCTCAAGGGAACCTTGCTCTCTCCA	372
Db	621	CCTGGCTGGACTTGGGGCTTGATGATCTCTGAAGCTCAAGGGAACCTTGCTCTCTCCA	680
Qy	373	GGCCTTCTCCAAAGACAGTTTTCCAAAACTGAGGATCCTGGAGTGCTTATGCTACCA	432
Db	681	GGCCTTCTCCAAAGACAGTTTTCCAAAACTGAGGATCCTGGAGTGCTTATGCTACCA	740
Qy	433	GTGCTGCTCCATGGGATGTGTGCAGCTTCTCAAGGCTCTGGGCACTGGGAGGCTGA	492
Db	741	GTGCTGCTCCATGGGATGTGTGCAGCTTCTCAAGGCTCTGGGCACTGGGAGGCTGA	800
Qy	493	AGACCTTCACCTTCATGATGAGGAGTCTTCAAAAGAGCCCTGGGCTCTCTGGCAGACA	552
Db	801	AGACCTTCACCTTCATGATGAGGAGTCTTCAAAAGAGCCCTGGGCTCTCTGGCAGACA	860
Qy	553	AGCAGAGAACCACTATGACACGAGACCTGGATGAGCTCCAGCTGAGATGGAGGACTCAAA	612
Db	861	AGCAGAGAACCACTATGACACGAGACCTGGATGAGCTCCAGCTGAGATGGAGGACTCAAA	920
Qy	613	GCACACCCAGTGTCCAGTGTAGCCCTATCTCCAGGCCCTTCAAGCCCTGTGAGTACCT	672
Db	921	GCACACCCAGTGTCCAGTGTAGCCCTATCTCCAGGCCCTTCAAGCCCTGTGAGTACCT	980
Qy	673	CTTTGAAAGCTGGGGCATCCGCCCTGGCGTGGCGGCATCGTGTGCTCTCCGTGCTCTG	732
Db	981	CTTTGAAAGCTGGGGCATCCGCCCTGGCGTGGCGGCATCGTGTGCTCTCCGTGCTCTG	1040
Qy	733	CAATGGACTGGTGTGCTGACCGGTTCGCTGGCGGCCTGCCGCCCTTCCCGGCTCAA	792
Db	1041	CAATGGACTGGTGTGCTGACCGGTTCGCTGGCGGCCTGCCGCCCTTCCCGGCTCAA	1100
Qy	793	GTTTGTGGTGGTGGCATTTGACGGGCCAACACCTTGACTGGGCATTTCTGTGGCCCTTCT	852
Db	1101	GTTTGTGGTGGTGGCATTTGACGGGCCAACACCTTGACTGGGCATTTCTGTGGCCCTTCT	1160
Qy	853	AGCCTCAGTCAGTCCCTGACTTTGTTGTCAGTTCCTGAGTACGAGGCCGCTGGGAGAC	912
Db	1161	AGCCTCAGTCAGTCCCTGACTTTGTTGTCAGTTCCTGAGTACGAGGCCGCTGGGAGAC	1220
Qy	913	GGGCTAGGCTGCGGGCCCACTGGCTTCTCGGCACTACTTGGGTCCGAGGAGCATCGGTGCT	972
Db	1221	GGGCTAGGCTGCGGGCCCACTGGCTTCTCGGCACTACTTGGGTCCGAGGAGCATCGGTGCT	1280
Qy	973	GCTGCTCACTTGGCCGAGTGGAGGCTCTCGGTCTCTGCTGTCTGCTGGGCTATGG	1032
Db	1281	GCTGCTCACTTGGCCGAGTGGAGGCTCTCGGTCTCTGCTGTCTGCTGGGCTATGG	1340

Qy	1033	GAACTCCCCCTCCTGGCGACGTTTCAGCAGGGGCTCTAGGCTGCCCTGGCACTGGCAGG	1092
Db	1341	GAACTCCCCCTCCTGGCGACGTTTCAGCAGGGGCTCTAGGCTGCCCTGGCACTGGCAGG	1400
Qy	1093	GCTGGCCGCGCAGCTGCCCCCTGGCCTCAGTGGGAGAAATACGGGGCCTCCCCACCTCTGCCT	1152
Db	1401	GCTGGCCGCGCAGCTGCCCCCTGGCCTCAGTGGGAGAAATACGGGGCCTCCCCACCTCTGCCT	1460
Qy	1153	GCCTACGCGCCACCTGAGGGTCAGCCAGCAGCCCTGGGCTTACCCGTGGCCCTGGTGAT	1212
Db	1461	GCCTACGCGCCACCTGAGGGTCAGCCAGCAGCCCTGGGCTTACCCGTGGCCCTGGTGAT	1520
Qy	1213	GATGAATCTCTCTGTTCTGTTCTGGTGGCGGCTTACATCAAACTGACTGTGACCT	1272
Db	1521	GATGAATCTCTCTGTTCTGTTCTGGTGGCGGCTTACATCAAACTGACTGTGACCT	1580
Qy	1273	GC CGCGGGCGACTTTGAGCCCTGTGGGACTGCGCCATGGTGGAGCACTGGGCTGGCT	1332
Db	1581	GC CGCGGGCGACTTTGAGCCCTGTGGGACTGCGCCATGGTGGAGCACTGGGCTGGCT	1640
Qy	1333	CATCTTCGAGAGAGGGCTCTCTACTGTCCCGTGGCCTTCTCAGCTTCGCTCCATGCT	1392
Db	1641	CATCTTCGAGAGAGGGCTCTCTACTGTCCCGTGGCCTTCTCAGCTTCGCTCCATGCT	1700
Qy	1393	GGGCTCTTCCCTGTCAACGCCCGAGGCCGTCAAGTCTGTCTGCTGTGGTGTGCTGCCCT	1452
Db	1701	GGGCTCTTCCCTGTCAACGCCCGAGGCCGTCAAGTCTGTCTGCTGTGGTGTGCTGCCCT	1760
Qy	1453	GCTGCTGTGCTCAACCCACTGTGTACTGTCTTCAACCCCACTTCGCGGAGTACCT	1512
Db	1761	GCTGCTGTGCTCAACCCACTGTGTACTGTCTTCAACCCCACTTCGCGGAGTACCT	1820
Qy	1513	TCCGCGGCTTCCGCGCCCGCGAGGAGCTCAGGGCCCTAGCCTATGCTCGGCGCGGGA	1572
Db	1821	TCGGCGGCTTCCGCGCCCGCGAGGAGCTCAGGGCCCTAGCCTATGCTCGGCGCGGGA	1880
Qy	1573	GCTGGAGAGAGCTCTCTGTGATTCTACCCAGGCGCTGGTAGCCTTCTCTGATGTGATCT	1632
Db	1881	GCTGGAGAGAGCTCTCTGTGATTCTACCCAGGCGCTGGTAGCCTTCTCTGATGTGATCT	1940
Qy	1633	CATTCTGGAAGCTTCTGAACTGGGGGCCCTTGGGCTGGAGACCTATGGCTTCCCTC	1692
Db	1941	CATTCTGGAAGCTTCTGAACTGGGGGCCCTTGGGCTGGAGACCTATGGCTTCCCTC	2000
Qy	1693	AGTGAACCTCATCTCTGTGTCAGCAGCCAGGGGCCCGCCAGGCTGGAGGCGACCACTTGT	1752
Db	2001	AGTGAACCTCATCTCTGTGTCAGCAGCCAGGGGCCCGCCAGGCTGGAGGCGACCACTTGT	2060
Qy	1753	AGAGCCAGAGGGGAAACCACTTTTGGGAAACCCCAACCCCTCCATGGATGGAGAACTGTGCT	1812
Db	2061	AGAGCCAGAGGGGAAACCACTTTTGGGAAACCCCAACCCCTCCATGGATGGAGAACTGTGCT	2120
Qy	1813	GAGGCGAGAGGGATCTACGCCAGCAGGTGGAGCTTGTTCAGGGGGTGGCGGCTTTCAGCC	1872
Db	2121	GAGGCGAGAGGGATCTACGCCAGCAGGTGGAGCTTGTTCAGGGGGTGGCGGCTTTCAGCC	2180
Qy	1873	CTCTGGCTGGCCCTTGTCTTCAACGTGTAAATATCCCTCCCACTTCTTCTTCCCTC	1932
Db	2181	CTCTGGCTGGCCCTTGTCTTCAACGTGTAAATATCCCTCCCACTTCTTCTTCCCTC	2240
Qy	1933	TCCTTCCCTTCTCTCTCCCTTGGTGAATGATGGCTGTCTTAAACAAATAACAAACCA	1992
Db	2241	TCCTTCCCTTCTCTCTCCCTTGGTGAATGATGGCTGTCTTAAACAAATAACAAACCA	2300
Qy	1993	AAACTCAGCAGTGTGATCTATAGCAGATGGCCAGTACCTGGCTCACCTGATCACTCT	2052
Db	2301	AAACTCAGCAGTGTGATCTATAGCAGATGGCCAGTACCTGGCTCACCTGATCACTCT	2360
Qy	2053	CTCCTGTGACCATCACCAAGGGTGTCTTGGCTGGCTTCCCTTGGCTTCCCTCAGC	2112
Db	2361	CTCCTGTGACCATCACCAAGGGTGTCTTGGCTGGCTTCCCTTGGCTTCCCTCAGC	2420
Qy	2113	TTCACTTGTATCTGGGCTCTTCTTGTCTGTAAGCTGTGGAGCTGGAGACCTGGA	2172

Db 2421 TTCACTTTGATCTGGGCTCTCTCTGTCATGCTGAAGCTGTGGACACAGACCTGGA 2480
Qy 2173 CTTTGTCTGCTTAAGGGAATAGAGGAAGTAAAGACAGTGAAGGGTGGAGGTTGATC 2232
Db 2481 CTTTGTCTGCTTAAGGGAATAGAGGAAGTAAAGACAGTGAAGGGTGGAGGTTGATC 2540
Qy 2233 AGGGCAGTGGACAGGAGACCTCACAGAGAAAGGCTGGAAGGTGATTCCCGTGTGA 2292
Db 2541 AGGGCAGTGGACAGGAGACCTCACAGAGAAAGGCTGGAAGGTGATTCCCGTGTGA 2600
Qy 2293 CTCATGGATAGATACAAAGTGTCTTCATGTACCATTAATCTTGACATATGCCATGCAT 2352
Db 2601 CTCATGGATAGATACAAAGTGTCTTCATGTACCATTAATCTTGACATATGCCATGCAT 2660
Qy 2353 AAAGACTTCTTATTAATAAGCTTTGGAAGAGATTAAATAAAAAAAAAAAAAA 2403
Db 2661 AAAGACTTCTTATTAATAAGCTTTGGAAGAGATTAAATAAAAAAAAAAAAAA 2711

RESULT 5
BC047905
LOCUS
DEFINITION Homo sapiens leucine-rich repeat-containing G protein-coupled
ACCESSION BC047905
VERSION BC047905.1 GI:29126870
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens

REFERENCE
AUTHORS

1 (bases 1 to 3286)
Strauberg, R.L., Feingold, E.A., Grouse, L.H., Derge, J.G.,
Klausner, R.D., Collins, F.S., Wagner, L., Shenmen, C.M., Schuler, G.D.,
Altschul, S.F., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Bhat, N.K.,
Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J.J., Hsieh, F.,
Diatchenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L.,
Stapleton, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L.,
Scheetz, T.E., Brownstein, M.J., Ustin, T.B., Toohiyuki, S.,
Carninci, P., Prange, C., Raha, S.S., Loquellano, N.A., Peters, G.J.,
Abramson, R.D., Mullany, S.J., Bosak, S.A., McEwan, P.J.,
McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S.,
Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hulyk, S.W.,
Villalón, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A.,
Fahey, J., Helton, E., Kettman, M., Madan, A., Rodriguez, S.,
Sanchez, A., Whiting, M., Madan, A., Young, A.C., Shevchenko, Y.,
Bouffard, G.G., Blakesley, R.W., Touchman, J.W., Green, E.D.,
Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M.,
Butterfield, Y.S., Krzywinski, M.I., Skalska, U., Smailus, D.E.,
Schnerch, A., Schein, J.E., Jones, S.J., and Marra, M.A.
Generation and initial analysis of more than 15,000 full-length
human and mouse cDNA sequences

JOURNAL
MEDLINE
PUBMED
REFERENCE
AUTHORS
TITLE
JOURNAL

Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
22388257
12477932
2 (bases 1 to 3286)
Strauberg, R.
Direct Submission
Submitted (03-MAR-2003) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA
NIH-MGC Project URL: <http://mgc.nci.nih.gov>
Contact: MGC help desk
Email: cgapsb-remail.nih.gov
Tissue Procurement: Life Technologies, Inc.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (ILNL)
DNA Sequencing by: Sequencing Group at the Stanford Human Genome
Center, Stanford University School of Medicine, Stanford, CA 94305
Web site: <http://www-shgc.stanford.edu>
Contact: (Dickson, Mark) mcd@paxil.stanford.edu

Dickson, M., Schmutz, J., Grimwood, J., Rodriguez, A., and Myers, R. M.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/ILNL at: <http://image.llnl.gov>
Series: IRAC Plate: 107 Row: d Column: 11.

FEATURES
source
1. 3286
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5212508"
/tissue_type="Blood, adult leukocytes"
/clone_lib="NIH MGC_118"
/lab_host="DH10B"
/note="Vector: pCMV-SPORT6"
<1..3286
/gene="LGR6"
/note="synonyms: FLJ14471, VTS20631"
/db_xref="LocusID:59352"
/db_xref="MIM:606653"
<1..2774
/gene="LGR6"
/codon_start=3
/product="LGR6 protein"
/protein_id="AAH47905.1"
/db_xref="GI:29126871"
/db_xref="LocusID:59352"
/db_xref="MIM:606653"
/translation="KGKPTGGMRPRLLTVQVSIISARDLSMNNLTLPGLFHH
LRFLEELRSGNHLSHIPGAFSLYSLKILMLQNQLGGIPAAELWELPSQSRLD
ANLISLVPERSFEGLSLRLHMLDDNALTEIPVALNNLPALQMTALNLRISHIPDY
AFQNLTSVLVHLHNNRIQLQTHSFEGLHNLTLDLNLYNKLOPFPVAIRTLGRLOEL
GFNNNTKAIPKAFMGNPILLOTIFYDNPQIFVGRSAFOYLPKLTLSLANGAMDOIE
FPDLKGTSLLEILTTRAGIRLLPSGMCQQLPRLRLVLELSHNOIEELPSLRCOKLEE
IPLOHNRWEIGATDQLSLSQALDLSWNAIRSIHPEAFYATLHSLVLDLDTDLNLT
LPLAGLGLMLKLGKLNLSQAFSDSPFKRLILEVPAYQCCPYGMCASFDFKASGQ
WEADHLDDDESSKRLGLARQAHENYDQDLQLEMDSPKHPVSVQSPPTGPF
KPEYLFELSGWIRLAVWAIIVLLSVLCNGLVLLTVFAGGPVLPVPKVVVGAAGANTL
TGISGILLASVADLTFGQSEYGARWETGLGRATGLAVLGASVSLTLLAAVQCS
VSVSCVRYGKSPSLGVRAGVGLCLALAGLAALPLASVGEYGNASPLCLYAPPEGQ
PAALGTVLVLMNNSFCFLVAGAYIKLYCDLPGRDFEAVWDCMVRHVAWLIFADGL
LYCPVAFSLFASMLGLFPVTPAEVKSLLVLPACNPLLLYLLPHFDDLRLLR
PRAGSGFLAYAAAGLEKSDSTQALVAFSDVDLILEASEAGRPGLFTYGPFSVT
LISQOCPARLEGHSGVPEBGNHFGNPQSPMDCELLLRABGSTPAGGSLGGGFGFP
SGLAFASHV"

CDS
Query Match 90.4%; Score 2247.8; DB 9; Length 3286;
Best Local Similarity 99.6%; Pred. No. 0;
Matches 2263; Conservative 0; Mismatches 7; Indels 1; Gaps 1;
Qy 133 AGCGGCTCCAAACACACCGCATCTGGAAATTTGGAGTNGACACCTTCAGCCAGCTGAG 192
Db 1005 ATCGGCTCCAAACACACCGCATCTGGAAATTTGGAGT -GACACCTTCAGCCAGCTGAG 1063
Qy 193 CTCCTGCAAGCCCTGGATCTTAGCTGAAGCCATCCGGTCCATCCACCTCAGGCGCTT 252
Db 1064 CTCCTGCAAGCCCTGGATCTTAGCTGAAGCCATCCGGTCCATCCACCTCAGGCGCTT 1123
Qy 253 CTCACCCCTGCATCTCCCTGGTCAAGCTGGAGCTGACAGACACCAAGCTGACACACTGCC 312
Db 1124 CTCACCCCTGCATCTCCCTGGTCAAGCTGGAGCTGACAGACACCAAGCTGACACACTGCC 1183
Qy 313 CTCCTGCGACTTGGGGCTTGATGCATCTGAAGCTCAAGGGAACCTTGTCTCTCCCA 372
Db 1184 CTCCTGCGACTTGGGGCTTGATGCATCTGAAGCTCAAGGGAACCTTGTCTCTCCCA 1243
Qy 373 GGCCTTCTCCAAAGGACAGTTTCCCAAACTCAGGATCTCGAGGTGCTTATGCTACCA 432
Db 1244 GGCCTTCTCCAAAGGACAGTTTCCCAAACTCAGGATCTCGAGGTGCTTATGCTACCA 1303
Qy 433 GTGCTGTCCCTATGGGATGTGTGCCAGCTTCTTCAAGGCTCTCGGCGAGTGGAGGCTGA 492

ORIGIN

Db	1304	 GTGCTGTCCTTATGGGATGTGTGCAGCTTCTTCAAGGCCCTCTGGGCAGTGGGAGGCTGA	1363
Qy	493	AGACCTTCACCTTGATGATGAGGAGTCTTTCAAAAAGGCCCTGGCCCTCTTCCTGCGAGACA	552
Db	1364	AGACCTTCACCTTGATGATGAGGAGTCTTCAAAAAGGCCCTGGCCCTCTTCCTGCGAGACA	1423
Qy	553	AGCAGAGAACCACTATGACCAAGGACCTGGATGAGCTCCAGCTGGAGATGGAGGACTCAAA	612
Db	1424	AGCAGAGAACCACTATGACCAAGGACCTGGATGAGCTCCAGCTGGAGATGGAGGACTCAAA	1483
Qy	613	GCCACACCCAGTGTCCAGTGTAGCCCTACTTCCAGGCCCTTCAAGGCCCTGTGAGTACCT	672
Db	1484	GCCACACCCAGTGTCCAGTGTAGCCCTACTTCCAGGCCCTTCAAGGCCCTGTGAGTACCT	1543
Qy	673	CTTTGAAAGCTGGGGCACTCCGCTGGCCGTGTGGGCCCATCGTGTGCTCTCCGTGCTCTG	732
Db	1544	CTTTGAAAGCTGGGGCATCCGCCGTGTGGGCCCATCGTGTGCTCTCCGTGCTCTG	1603
Qy	733	CAATGGGACTGTGTCTGTGACCGCTTTGCGCTGGCGGCCCTGCCCCCTGCCCCCGGTCAA	792
Db	1604	CAATGGGACTGTGTCTGTGACCGCTTTGCGCTGGCGGCCCTGTCCCTGCCCCCGGTCAA	1663
Qy	793	GTTTGTGTAGGTGGGATTTGAGGGGCCAAACACCTTGACTGGCAATTCCTGTGGCCTTCT	852
Db	1664	GTTTGTGTAGGTGGGATTTGAGGGGCCAAACACCTTGACTGGCAATTCCTGTGGCCTTCT	1723
Qy	853	AGCCTCAGTGCATGCCCTGACCTTTTGTCAGTTCCTGAGTACGGAGCCCGCTGGGAGAC	912
Db	1724	AGCCTCAGTGCATGCCCTGACCTTTTGTCAGTTCCTGAGTACGGAGCCCGCTGGGAGAC	1783
Qy	913	GGGGCTAGGCTGCCGGGCCACTGGCTTCTCGGCAGTACTTGGGTGCGAGGCATCGGTGCT	972
Db	1784	GGGGCTAGGCTGCCGGGCCACTGGCTTCTCGGCAGTACTTGGGTGCGAGGCATCGGTGCT	1843
Qy	973	GCTGCTCACTGTGCGCGAGTGCAGTGCAGCGTTCCTGCTTCTCCTGTGTGCGGGCCTATGG	1032
Db	1844	GCTGCTCACTGTGCGCGAGTGCAGTGCAGCGTTCCTGCTTCTCCTGTGTGCGGGCCTATGG	1903
Qy	1033	GAACTCCCCCTCCCTGGGACGTTTCAGACGAGGGTCTTAGGCTGCTGCGACCTGGCAGG	1092
Db	1904	GAACTCCCCCTCCCTGGGACGTTTCAGACGAGGGTCTTAGGCTGCTGCGACCTGGCAGG	1963
Qy	1093	GCTGGCCGCGCACTGCCCTTGGGCTCAGTGGGAGAAATACGGGGCTTCCCCACTCTGCCT	1152
Db	1964	GCTGGCCGCGCGTGTCCCCCTGGGCTCAGTGGGAGAAATACGGGGCTTCCCCACTCTGCCT	2023
Qy	1153	GCCCTACGCGCACCTGAGGGTACGACGAGACCCCTGGGGTTACCGTGGCCCTGGTGTAT	1212
Db	2024	GCCCTACGCGCACCTGAGGGTACGACGAGACCCCTGGGGTTACCGTGGCCCTGGTGTAT	2083
Qy	1213	GATGAACTCCTTCTGTTTCTGTGTCGTGGCGGTGCCTACATCAAACTGTACTGTGACCT	1272
Db	2084	GATGAACTCCTTCTGTTTCTGTGTCGTGGCGGTGCCTACATCAAACTGTACTGTGACCT	2143
Qy	1273	GCCGCGGGCGACTTTGAGGCCGTGTGGGACTGCGCATGGTGAAGGCACGTGGGCTCGCT	1332
Db	2144	GCCGCGGGCGACTTTGAGGCCGTGTGGGACTGCGCCATGGTGAAGGCACGTGGGCTCGCT	2203
Qy	1333	CATCTTCGACAGCGGCTCTCTACTGTCCCGTGGCCCTTCCCTCAGCTTCGCCCTCCATGCT	1392
Db	2204	CATCTTCGACAGCGGCTCTCTACTGTCCCGTGGCCCTTCCCTCAGCTTCGCCCTCCATGCT	2263
Qy	1393	GGGCTCTTCCCTGTACGCCCGAGGCGGTCAAGTCTGTCTGTGTGGTGTGTCGCCCT	1452
Db	2264	GGGCTCTTCCCTGTACGCCCGAGGCGGTCAAGTCTGTCTGTGTGGTGTGTCGCCCT	2323
Qy	1453	GCCTGCTGCTCAAACTGCTGTACTGCTCTTCAACCCCACTTCGGGATGACCT	1512
Db	2324	GCCTGCTGCTCAAACTGCTGTACTGCTCTTCAACCCCACTTCGGGATGACCT	2383
Qy	1513	TCGGCGGCTTCGGGCCCGCGAGGGGACTCAGGGCCCTCAGCTTATGCTGCGGCGGGGA	1572

Db	2384	TCGGCGGCTTCCGGCCCCCGCAGAGGGGACTCAGAGGCCCTTAGCTATGCTTGCGGCCGGGGGA	2443
Qy	1573	GCTGGAGAAGAGCTCTCTGTGATTCTACCCAGGGCCCTGGTAGCCCTTCTCTGATGTGGATCT	1632
Db	2444	GCTGGAGAAGAGCTCTCTGTGATTCTACCCAGGGCCCTGGTAGCCCTTCTCTGATGTGGATCT	2503
Qy	1633	CATTCTGGAAAGCTTTCTGAAGCTGGCGGGCCCCCTTGGGCTCGAGACCCTATAGGCTTCCCCTC	1692
Db	2504	CATTCTGGAAAGCTTTCTGAAGCTGGCGGGCCCCCTTGGGCTCGAGACCCTATAGGCTTCCCCTC	2563
Qy	1693	AGTGACCCCTCATCTCTGTCTCAGCAGCAGAGGGGCCCCAGAGCTGGAGGGCAGCCAATTCTGT	1752
Db	2564	AGTGACCCCTCATCTCTGTCTCAGCAGCAGAGGGGCCCCAGAGCTGGAGGGCAGCCAATTCTGT	2623
Qy	1753	AGAGCCAGAGGGGAACCACTTTTGGGAACCCCAACCTCCATGGATGGAGAACCTGCCTGT	1812
Db	2624	AGAGCCAGAGGGGAACCACTTTTGGGAACCCCAACCTCCATGGATGGAGAACCTGCCTGT	2683
Qy	1813	GAGGCGAAGAGGATCTACGCCAGCAGGCTGGTGAAGGCTTGTTCAGGGGGTGGCGGCTTTCAGCC	1872
Db	2684	GAGGCGAAGAGGATCTACGCCAGCAGGCTGGTGAAGGCTTGTTCAGGGGGTGGCGGCTTTCAGCC	2743
Qy	1973	CTCTGGCTTGGCCCTTTGCTTTCACACGTGTAAATATCCCTCCCACTTCTCTCTCCCTC	1932
Db	2744	CTCTGGCTTGGCCCTTTGCTTTCACACGTGTAAATATCCCTCCCACTTCTCTCTCCCTC	2803
Qy	1933	TCTTCCCTTTCTCTCTCCCTCCGGTGAATGATGGCTGCTTCTATAAAACAATAACAACCA	1992
Db	2804	TCTTCCCTTTCTCTCTCCCTCCGGTGAATGATGGCTGCTTCTATAAAACAATAACAACCA	2863
Qy	1993	AAACTCAGCAGTGTGATCTATAGCAGGATGGCCCACTGATCTGGCTCCACTGATCACCTCT	2052
Db	2864	AAACTCAGCAGTGTGATCTATAGCAGGATGGCCCACTGATCTGGCTCCACTGATCACCTCT	2923
Qy	2053	CTCTGTGACCATCAACAACGGGTGCTCTTGGCTGGCTTTCCTTGGCTTCTCTCAGC	2112
Db	2924	CTCTGTGACCATCAACAACGGGTGCTCTTGGCTGGCTTTCCTTGGCTTCTCTCAGC	2983
Qy	2113	TTCACTTTGATACTGGGCTCTTCTTGTCTCATGTCTGAAGCTGTGGACCAAGAGACCTGGA	2172
Db	2984	TTCACTTTGATACTGGGCTCTTCTTGTCTCATGTCTGAAGCTGTGGACCAAGAGACCTGGA	3043
Qy	2173	CTTTTGTCTCTTAAGGGAATGAGGGAAGTAAAGACAGTGAAGGGGTGAGGGTTGATC	2232
Db	3044	CTTTTGTCTCTTAAGGGAATGAGGGAAGTAAAGACAGTGAAGGGGTGAGGGTTGATC	3103
Qy	2233	AGGCGACAGTGGACAGGAGACCTCACAGAGAAAGGCTGGAGAGGTGATTTCCGTGTGA	2292
Db	3104	AGGCGACAGTGGACAGGAGACCTCACAGAGAAAGGCTGGAGAGGTGATTTCCGTGTGA	3163
Qy	2293	CTCATGGATAGGATACAAAATGTTCATGTACCATTAATCTTTGACATATGCCATGCAT	2352
Db	3164	CTCATGGATAGGATACAAAATGTTCATGTACCATTAATCTTTGACATATGCCATGCAT	3223
Qy	2353	AAAGACTTCTCTATAAATAAGCTTTTGGAGAGATTAAAAAIAAAAAAAAAA	2403
Db	3224	AAAGACTTCTCTATAAATAAGCTTTTGGAGAGATTAAAAAIAAAAAAAAAA	3274
RESULT 6			
CQ840811			
LOCUS	CQ840811 3306 bp DNA linear PAT 29-JUL-2004		
DEFINITION	Sequence 1 from Patent WO2004058818.		
ACCESSION	CQ840811		
VERSION	CQ840811.1 GI:50838402		
KEYWORDS	Homo sapiens (human)		
SOURCE	Homo sapiens		
ORGANISM	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.		
REFERENCE	1		
AUTHORS	Muda, M.		
TITLE	Spliced variants of lgr6		

JOURNAL

Patent: WO 2004058818-A 1 15-JUL-2004;

Applied Research Systems ARS Holding N.V. (AN)

FEATURES

Location/Qualifiers

source

1..3306

/organism="Homo sapiens"

/mol_type="unassigned DNA"

/db_xref="taxon:9606"

47..2794

/note="unnamed protein product"

/codon_start=1

/protein_id="CAH05693.1"

/db_xref="GI:50838403"

/translation="MGRRLTLVCOVSIISARDLSMNLTELQELPHHLRFLBELR
LSGNHLSHPGAFSGLYSLKILMIONNQLGIPABALWELPQLSRLDANLISLVP
ERSFGLSLRHLRDDNALTEIPVRLNLPALQAMTLALNRISHIEDYAFQNLISL
VYLHNHNRIOHLGTHSPENHLETLNLYNKLQEPFVAIRTLRLQELGPFHNNIK
AIPKAFMNRIPLOTHIFPDNPQIFVGRSAFOYLPLKLTLSLNGAMDIOEPDLKGT
SLEILTFRAGIRLLPSMCOQLPRLRVLELSHNOIELPSLRQCKLEIEIGLOHNR
WEIGADTFSQSSLOALDLSNARSIHPEAFSTLSHLVKLDLDTLQITLPLAGLGG
LMHLKLNGLALSQAFSKDPSFPKRLILEVPYAYQCCPYGMCASFPAKSGWEADLHG
DDEBSKRLPLGLLRQAEHNYDDLDLEQLEMEDSKPHSPQCSPTPGFPKPEYLFE
SWGIRLAWAIVLLSVLNGVLVLTVPFAGGPVLPVVFVVGATAGANTLTGISCLL
YKSPSLGSVRAGVLGCLAGLAAALPLASVGEXGASPLCLPYAPPGQPAALGFTV
ALVMNSFCFLVAGAYIKYCDLPRGDFEAVMDCAVRHVAWLIFADGLLYCPVAFI
SPASMLGFPVTPPEAKSVLLVLPPLPACINPLLYLLFNPHFRDLRLRPRAGDSGP
LAYAAGLELSSCDSTQALVAFSDVLDILEASEAGRPGLFTYGFPSVTLISQOPG
APRLGSHCVPEGNHFNQPPQPSMDGELLRLRAEGSTPAGGGLSGGGGFPQPSGLAFASH
V"

ORIGIN

Query Match 90.4%; Score 2247.8; DB 6; Length 3306;

Best Local Similarity 99.6%; Pred. No. 0;

Matches 2263; Conservative 0; Mismatches 7; Indels 1; Gaps 1;

Db 1564 CTTTGAAGCTGGGGGATCCGCTGGCGTGTGGGCCATCGTGTGTCTCTCGTCTCTG 1623
Qy 733 CAATGGACTGGTGTCTGCTGACCGTGTTCGCTGGGGGGCTGCCCCCTGCCCCCGTCAA 792
Db 1624 CAATGGACTGGTGTCTGCTGACCGTGTTCGCTGGGGGGCTGCCCCCTGCCCCCGTCAA 1683
Qy 793 GTTTGTGTAGTGGGATGTGAGCGGCCCAACACCTTGTGACTGGCATTTCTGTGGCTTCT 852
Db 1684 GTTTGTGTAGTGGGATGTGAGCGGCCCAACACCTTGTGACTGGCATTTCTGTGGCTTCT 1743
Qy 853 AGCCTCAGTCGATGCCCTGACCTTTGGTTCAGTCTCTCTGAGTACGGAGCCCGTGGGAGAC 912
Db 1744 AGCCTCAGTCGATGCCCTGACCTTTGGTTCAGTCTCTCTGAGTACGGAGCCCGTGGGAGAC 1803
Qy 913 GGGGTAGGCTGCGGGGCCACCTGGCTTCTCTGGCAGTACTTGGGTGGAGGATCGGTGCT 972
Db 1804 GGGGTAGGCTGCGGGGCCACCTGGCTTCTCTGGCAGTACTTGGGTGGAGGATCGGTGCT 1863
Qy 973 GCTGCTCACTCTGGCGGAGTGCAGTGAGCGTCTCCGTCTCTGTGTCCGGGCCATATGG 1032
Db 1864 GCTGCTCACTCTGGCGGAGTGCAGTGAGCGTCTCCGTCTCTGTGTCCGGGCCATATGG 1923
Qy 1033 GAAGTCCCTCCCTGGGACGCTTCGAGCAGGGGTCTTAGGCTGCCTGGCACATGGCAGG 1092
Db 1924 GAAGTCCCTCCCTGGGACGCTTCGAGCAGGGGTCTTAGGCTGCCTGGCACATGGCAGG 1983
Qy 1093 GCTGGCCCGCCACCTGCCCTTGGGCTCAGTGGGAGAAATACGGGGCTCCCCACTCTGCT 1152
Db 1984 GCTGGCCCGCCGCTGCCCTTGGGCTCAGTGGGAGAAATACGGGGCTCCCCACTCTGCT 2043
Qy 1153 GCCCTAGCGCCACCTGAGGCTCAGCAGAGCCCTGGGCTTACCGTGGGCCCTGGTGTAT 1212
Db 2044 GCCCTAGCGCCACCTGAGGCTCAGCAGAGCCCTGGGCTTACCGTGGGCCCTGGTGTAT 2103
Qy 1213 GATGAATCTCTCTGTTCTGCTGGCGCGTGCCTTACATCAAACTGTACTGTGACT 1272
Db 2104 GATGAATCTCTCTGTTCTGCTGGCGCGTGCCTTACATCAAACTGTACTGTGACT 2163
Qy 1273 GCCGCGGGCCACCTTTGAGGCGCTGTGGGACTGGGCCATGTGTAGGACAGTGGCTGGCT 1332
Db 2164 GCCGCGGGCCACCTTTGAGGCGCTGTGGGACTGGGCCATGTGTAGGACAGTGGCTGGCT 2223
Qy 1333 CATCTTGCAGACGGGCTCTCTACTGTCCTGGGCTTCTCAGCTTTCAGCTTTCATGCT 1392
Db 2224 CATCTTGCAGACGGGCTCTCTACTGTCCTGGGCTTCTCAGCTTTCATGCTTTCATGCT 2283
Qy 1393 GGGCTCTTCCCTCTCAGCGCGGAGCGCTCAAGTCTGCTGCTGTGGTGTGCTGCCCT 1452
Db 2284 GGGCTCTTCCCTCTCAGCGCGGAGCGCTCAAGTCTGCTGCTGTGGTGTGCTGCCCT 2343
Qy 1453 GCCTGCTGCTCAACCCACCTGCTGCTCTTCAACCCCACTTCCGGGATGACT 1512
Db 2344 GCCTGCTGCTCAACCCACCTGCTGCTCTTCAACCCCACTTCCGGGATGACT 2403
Qy 1513 TCGCGGGCTTCGGCCCCCGCAGGGGACTAGGGCCCCCTAGCCCTATGCTCGGGCGGGGA 1572
Db 2404 TCGCGGGCTTCGGCCCCCGCAGGGGACTAGGGCCCCCTAGCCCTATGCTCGGGCGGGGA 2463
Qy 1573 GCTGGAGAGAGCTCTGTGATTTCTACCCAGGCCCTGGTAGCTTCTGTGATGGATCT 1632
Db 2464 GCTGGAGAGAGCTCTGTGATTTCTACCCAGGCCCTGGTAGCTTCTGTGATGGATCT 2523
Qy 1633 CATCTGGAAGCTTCTGAAGCTGGCGGCCCTTGGGCTGGAGACCTATGCTTCCCCCTC 1692
Db 2524 CATCTGGAAGCTTCTGAAGCTGGCGGCCCTTGGGCTGGAGACCTATGCTTCCCCCTC 2583
Qy 1693 AGTGACCCCTCACTCTCTGTCAGCAGCCAGGGCCCCCAGGCTGGAGGAGCCATTGTGT 1752
Db 2584 AGTGACCCCTCACTCTCTGTCAGCAGCCAGGGCCCCCAGGCTGGAGGAGCCATTGTGT 2643
Qy 1753 AGAGCCAGAGGGGAAACCACTTTGGGAAACCCCAACCTTCCATGGATGGAGAACTGCTGCT 1812
Db 2644 AGAGCCAGAGGGGAAACCACTTTGGGAAACCCCAACCTTCCATGGATGGAGAACTGCTGCT 2703

JOURNAL

Patent: WO 2004058818-A 1 15-JUL-2004;

Applied Research Systems ARS Holding N.V. (AN)

FEATURES

Location/Qualifiers

source

1..3306

/organism="Homo sapiens"

/mol_type="unassigned DNA"

/db_xref="taxon:9606"

47..2794

/note="unnamed protein product"

/codon_start=1

/protein_id="CAH05693.1"

/db_xref="GI:50838403"

/translation="MGRRLTLVCOVSIISARDLSMNLTELQELPHHLRFLBELR
LSGNHLSHPGAFSGLYSLKILMIONNQLGIPABALWELPQLSRLDANLISLVP
ERSFGLSLRHLRDDNALTEIPVRLNLPALQAMTLALNRISHIEDYAFQNLISL
VYLHNHNRIOHLGTHSPENHLETLNLYNKLQEPFVAIRTLRLQELGPFHNNIK
AIPKAFMNRIPLOTHIFPDNPQIFVGRSAFOYLPLKLTLSLNGAMDIOEPDLKGT
SLEILTFRAGIRLLPSMCOQLPRLRVLELSHNOIELPSLRQCKLEIEIGLOHNR
WEIGADTFSQSSLOALDLSNARSIHPEAFSTLSHLVKLDLDTLQITLPLAGLGG
LMHLKLNGLALSQAFSKDPSFPKRLILEVPYAYQCCPYGMCASFPAKSGWEADLHG
DDEBSKRLPLGLLRQAEHNYDDLDLEQLEMEDSKPHSPQCSPTPGFPKPEYLFE
SWGIRLAWAIVLLSVLNGVLVLTVPFAGGPVLPVVFVVGATAGANTLTGISCLL
YKSPSLGSVRAGVLGCLAGLAAALPLASVGEXGASPLCLPYAPPGQPAALGFTV
ALVMNSFCFLVAGAYIKYCDLPRGDFEAVMDCAVRHVAWLIFADGLLYCPVAFI
SPASMLGFPVTPPEAKSVLLVLPPLPACINPLLYLLFNPHFRDLRLRPRAGDSGP
LAYAAGLELSSCDSTQALVAFSDVLDILEASEAGRPGLFTYGFPSVTLISQOPG
APRLGSHCVPEGNHFNQPPQPSMDGELLRLRAEGSTPAGGGLSGGGGFPQPSGLAFASH
V"

ORIGIN

Query Match 90.4%; Score 2247.8; DB 6; Length 3306;

Best Local Similarity 99.6%; Pred. No. 0;

Matches 2263; Conservative 0; Mismatches 7; Indels 1; Gaps 1;

Qy 133 AGCGCCTCCACACACCGCATCTGGGAATGGAGCTNGACACTTCAGCCAGCTGAG 192
Db 1025 ATCGGCCCTCCACACACCGCATCTGGGAATTTGGAGCT-GACACCTTCAGCCAGCTGAG 1083
Qy 193 CTCCCTGCAAGCCCTGGATCTTAGCTGGAACGCCATCCGGTCCATCCACCCTGAGGCCCT 252
Db 1084 CTCCCTGCAAGCCCTGGATCTTAGCTGGAACGCCATCCGGTCCATCCACCCTGAGGCCCT 1143
Qy 253 CTCACCTGCACTCCCTGGTCAAGCTGAGCTGACAGACACACAGCTGACACACTGCC 312
Db 1144 CTCACCTGCACTCCCTGGTCAAGCTGAGCTGAGCTGAGACACACAGCTGACACACTGCC 1203
Qy 313 CTTGGCTGCACTTGGGGCTTGATGATCATCTGAGCTCAAGGACCTTGGCTCTCCCA 372
Db 1204 CTTGGCTGCACTTGGGGCTTGATGATCATCTGAGCTCAAGGACCTTGGCTCTCCCA 1263
Qy 373 GGCTTCTCCAGGACAGTTTCCAAAACCTGAGGATCTGGAGTGCCCTTATGCCATCCA 432
Db 1264 GGCTTCTCCAGGACAGTTTCCAAAACCTGAGGATCTGGAGTGCCCTTATGCCATCCA 1323
Qy 433 GTGCTGCTCCCTATGGGATGTGTGCGAGCTTCTTCAAGGCCCTCTGGGAGTGGAGGCTGA 492
Db 1324 GTGCTGCTCCCTATGGGATGTGTGCGAGCTTCTTCAAGGCCCTCTGGGAGTGGAGGCTGA 1383
Qy 493 AGACCTTACCTTGATGATGAGGAGTCTTCAAAAAGGCCCTTGGGCCCTCTTGGCAGACA 552
Db 1384 AGACCTTACCTTGATGATGAGGAGTCTTCAAAAAGGCCCTTGGGCCCTCTTGGCAGACA 1443
Qy 553 AGCAGAGAACCACTATGACAGGACCTGGATGAGCTCCAGCTGGAGATGGAGGACTCAAA 612
Db 1444 AGCAGAGAACCACTATGACAGGACCTGGATGAGCTCCAGCTGGAGATGGAGGACTCAAA 1503
Qy 613 GCCACACCCAGTGTCCAGTGTAGCCCTACTCCAGGCCCTTCAAGCCCTGTGAGTACT 672
Db 1504 GCCACACCCAGTGTCCAGTGTAGCCCTACTCCAGGCCCTTCAAGCCCTGTGAGTACT 1563
Qy 673 CTTTGAAGCTGGGGCATCCGCTCGGCGTGTGGGCCATCGTGTGTCTCTCGTCTCTG 732

Matches 2260; Conservative 0; Mismatches 3; Indels 1; Gaps 1;			
Qy	133	AGCGCCTCCACACAAACCGCATCTGGGAATTTGGAGCTNGACACCTTCAGCCAGCTGAG	192
Dd	1063	ATCGGCCCTCCAAACACAAACCGCATCTGGGAATTTGGAGCT-GACACCTTCAGCCAGCTGAG	1121
Qy	193	CTCCCTGCAAGGCCCTTGGATCTTAGCTGGAAACGCCATCCGGTCCATCCACACCTGAGGCCCTT	252
Dd	1122	CTCCCTGCAAGGCCCTTGGATCTTAGCTGGAAACGCCATCCGGTCCATCCACACCTGAGGCCCTT	1181
Qy	253	CTCCACCTTGCACTCCCTGGTCAAGCTGGACCTGACAGAACCAACCAAGCTGACCAACATGCC	312
Dd	1182	CTCCACCTTGCACTCCCTGGTCAAGCTGGACCTGACAGAACCAACCAAGCTGACCAACATGCC	1241
Qy	313	CTGGCTGGACTTGGGGCTTGATGATCATCTGAAGCTCAAAGGAAACCTTGTCTCTCCCA	372
Dd	1242	CTGGCTGGACTTGGGGCTTGATGATCATCTGAAGCTCAAAGGAAACCTTGTCTCTCCCA	1301
Qy	373	GGCCTTCTCCAAAGGACAGTTTCCCAAAACTGAGGATCTGGAGGTGCCCTTATGCTACCA	432
Dd	1302	GGCCTTCTCCAAAGGACAGTTTCCCAAAACTGAGGATCTGGAGGTGCCCTTATGCTACCA	1361
Qy	433	GTGCTGCTCCCTATGGGATGTGTGCAGCTTCTTCAAGGCCCTCTGGGCAAGTGGGAGGCTGA	492
Dd	1362	GTGCTGCTCCCTATGGGATGTGTGCAGCTTCTTCAAGGCCCTCTGGGCAAGTGGGAGGCTGA	1421
Qy	493	AGACCTTCACTTGTATGATGAGGAGTCTTCAAAAGGCCCTTGGGCCCTCTTGGCAGACA	552
Dd	1422	AGACCTTCACTTGTATGATGAGGAGTCTTCAAAAGGCCCTTGGGCCCTCTTGGCAGACA	1481
Qy	553	AGCAGAGAACCACTATGACCAAGGACCTGGATGAGCTCCAGCTGGAGATGGAGGACTCAAA	612
Dd	1482	AGCAGAGAACCACTATGACCAAGGACCTGGATGAGCTCCAGCTGGAGATGGAGGACTCAAA	1541
Qy	613	GCCACACCCCACTAGTGTAGAGTGTAGCCCTACTCCAGGCCCTTCAAGGCCCTGTGAGTACT	672
Dd	1542	GCCACACCCCACTAGTGTAGAGTGTAGCCCTACTCCAGGCCCTTCAAGGCCCTGTGAGTACT	1601
Qy	673	CTTTGAAAGCTGGGGCATCCGCTGGCGGTGGGCCCATCTGTGTCTCTCCGTGCTCTG	732
Dd	1602	CTTTGAAAGCTGGGGCATCCGCTGGCGGTGGGCCCATCTGTGTCTCTCCGTGCTCTG	1661
Qy	733	CAATGGACTGTGTGCTGACCGTGTTCGCTGGGGGCTGCCGCCCTGCCCGCCGGTCAA	792
Dd	1662	CAATGGACTGTGTGCTGACCGTGTTCGCTGGGGGCTGCCGCCCTGCCCGCCGGTCAA	1721
Qy	793	GTTTGTGTAGGTGCGATPTGACGGCGCAACACCTTGACTGGCAATTTCTGTGCGCTTCT	852
Dd	1722	GTTTGTGTAGGTGCGATPTGACGGCGCAACACCTTGACTGGCAATTTCTGTGCGCTTCT	1781
Qy	853	AGCCTCAGTCAGTGCCTGACCTTTTGTGTCAGTTCTCTGAGTACGGAGCCGCTGGGAGAC	912
Dd	1782	AGCCTCAGTCAGTGCCTGACCTTTTGTGTCAGTTCTCTGAGTACGGAGCCGCTGGGAGAC	1841
Qy	913	GGGGTAGGCTGCGGGCCACTGTGCTTCTGGCAGTACTTGGGTGGAGGACATCGGTGCT	972
Dd	1842	GGGGTAGGCTGCGGGCCACTGTGCTTCTGGCAGTACTTGGGTGGAGGACATCGGTGCT	1901
Qy	973	GCTGCTCACTGTGGCGCAGTGCAGTGCAGCGTCTCCGTCTCTCTGTGTCCGGGCTATGG	1032
Dd	1902	GCTGCTCACTGTGGCGCAGTGCAGTGCAGCGTCTCCGTCTCTCTGTGTCCGGGCTATGG	1961
Qy	1033	GAAGTCCCTCTCCCTGGGACGGTTCGAGCAGGGGTCTTAGGCTGCCTGGACATGGCAGG	1092
Dd	1962	GAAGTCCCTCTCCCTGGGACGGTTCGAGCAGGGGTCTTAGGCTGCCTGGCAGTGGCAGG	2021
Qy	1093	GCTGGCGCGCACTGGCCCTGGCTCAGTGGGAGAAATACGGGGCTTCCCACTCTGCCT	1152
Dd	2022	GCTGGCGCGCACTGGCCCTGGCTCAGTGGGAGAAATACGGGGCTTCCCACTCTGCCT	2081
Qy	1153	GCCCTACGCGCCACTGAGGGTACGCCAGCAGCCCTGGGCTTACCGTGGGCCCTGGTGAT	1212
Dd	2082	GCCCTACGCGCCACTGAGGGTACGCCAGCAGCCCTGGGCTTACCGTGGGCCCTGGTGAT	2141

Qy	1213	GATGAACCTCCTTCTGTTTCTGTTGTCGTGGCGGTGCCTACATCAAACTGTACTGTGACCT	1272
Dd	2142	GATGAACCTCCTTCTGTTTCTGTTGTCGTGGCGGTGCCTACATCAAACTGTACTGTGACCT	2201
Qy	1273	GCCCGGGGGCGACTTTTGAGGCCGTGTGGGACTGCGCCATGATGAGGACGCTGGGCTGGCT	1332
Dd	2202	GCCCGGGGGCGACTTTTGAGGCCGTGTGGGACTGCGCCATGATGAGGACGCTGGGCTGGCT	2261
Qy	1333	CATCTTTCGAGAGAGGGGTCTCTACTGTCCCGTGGCTTCTCTCAGCTTTCGCTCCATGCT	1392
Dd	2262	CATCTTTCGAGAGAGGGGTCTCTACTGTCCCGTGGCTTCTCTCAGCTTTCGCTCCATGCT	2321
Qy	1393	GGGCTCTTTCCTGTACGCGCCGAGGGCGTCTAGTGTCTGCTGCTGCTGCTGCTGCTGCT	1452
Dd	2322	GGGCTCTTTCCTGTACGCGCCGAGGGCGTCTAGTGTCTGCTGCTGCTGCTGCTGCTGCT	2381
Qy	1453	GCCTGCTGCTCAACCCACTGTCTGTACTGTCTTCAACCCGCCCTTCCGGGATGACCT	1512
Dd	2382	GCCTGCTGCTCAACCCACTGTCTGTACTGTCTTCAACCCGCCCTTCCGGGATGACCT	2441
Qy	1513	TCGGCGGCTTCGGCCCGCGGAGGAGCTCAGGGCCCTTAGCCTATGCTGCGGCGGGGA	1572
Dd	2442	TCGGCGGCTTCGGCCCGCGGAGGAGCTCAGGGCCCTTAGCCTATGCTGCGGCGGGGA	2501
Qy	1573	GCTGGAGAGAGCTCTCTGTGATTTACTCCAGGGCCCTGGTAGCCTTCTCTGATGTGGATCT	1632
Dd	2502	GCTGGAGAGAGCTCTCTGTGATTTACTCCAGGGCCCTGGTAGCCTTCTCTGATGTGGATCT	2561
Qy	1633	CATTCTGGAAGCTTCTGAAGCTGGGGCGGCCCTTGGGCTGGAGACCTATGGCTTCCCTC	1692
Dd	2562	CATTCTGGAAGCTTCTGAAGCTGGGGCGGCCCTTGGGCTGGAGACCTATGGCTTCCCTC	2621
Qy	1693	AGTGACCTCATCTCTGTTCAGCAGCAGGGGGCCCCAGGCTGGAGGGCAGCCATTGTGT	1752
Dd	2622	AGTGACCTCATCTCTGTTCAGCAGCAGGGGGCCCCAGGCTGGAGGGCAGCCATTGTGT	2691
Qy	1753	AGAGCCAGAGGGGACCACTTTTGGAGACCCCAACCTTCCATGGATGGAGAACCTGTGCT	1812
Dd	2682	AGAGCCAGAGGGGACCACTTTTGGAGAACCCCAACCTTCCATGGATGGAGAACCTGTGCT	2741
Qy	1813	GAGGCGAGAGGGATCTACGCGCAGCAGGTGGAGGCTTGTGTCAGGGGGTGGCGGCTTCAGCC	1872
Dd	2742	GAGGCGAGAGGGATCTACGCGCAGCAGGTGGAGGCTTGTGTCAGGGGGTGGCGGCTTCAGCC	2801
Qy	1873	CTCTGGCTTGGCTTGTGCTTCAACAGGTGTAATAATCCCTCCCAATTTCTCTTCCCTC	1932
Dd	2802	CTCTGGCTTGGCTTGTGCTTCAACAGGTGTAATAATCCCTCCCAATTTCTCTTCCCTC	2861
Qy	1933	TCTTCCCTTCTCTCTCCCTCGGTGAATGATGGCTGCTTCTTAAACAAATACACCA	1992
Dd	2862	TCTTCCCTTCTCTCTCCCTCGGTGAATGATGGCTGCTTCTTAAACAAATACACCA	2921
Qy	1993	AAACTCAGCAGTGTGATCTATAGCAGGATGGCCAGTACCTGGCTCCACTGATCACCTCT	2052
Dd	2922	AAACTCAGCAGTGTGATCTATAGCAGGATGGCCAGTACCTGGCTCCACTGATCACCTCT	2981
Qy	2053	CTCTGTGACCATCAACACGGGTGCTCTTGGCTGGCTTTTCCCTTGGCTTCCCTCAGC	2112
Dd	2982	CTCTGTGACCATCAACACGGGTGCTCTTGGCTGGCTTTTCCCTTGGCTTCCCTCAGC	3041
Qy	2113	TTCACTTGTATCTGGGCTCTTCTTGTATCTCTGTAAGCTGTGGACAGAGACCTGGA	2172
Dd	3042	TTCACTTGTATCTGGGCTCTTCTTGTATCTCTGTAAGCTGTGGACAGAGACCTGGA	3101
Qy	2173	CTTTTGTCTCTTAAAGGAAATGAGGAAAGTAAAGACAGTGAAGGGGTGGAGGTTGATC	2232
Dd	3102	CTTTTGTCTCTTAAAGGAAATGAGGAAAGTAAAGACAGTGAAGGGGTGGAGGTTGATC	3161
Qy	2233	AGGGCAGTGGAGAGGGAGACCTCAAGAGAAAGGCTTGGAAAGGTGATTTCCCGTGTGA	2292
Dd	3162	AGGGCAGTGGAGAGGGAGACCTCAAGAGAAAGGCTTGGAAAGGTGATTTCCCGTGTGA	3221

QY 2293 CTCATGGATAGGATACAAATATGTTTCCATGTACCAATTAATCTTGACATATGCCATG 2352
Db 3222 CTCATGGATAGGATACAAATATGTTTCCATGTACCAATTAATCTTGACATATGCCATG 3281
QY 2353 AAGAGCTTCCCTATTAAATTAAGCTTTGGAGAGATTAATAAAAAA 2396
Db 3282 AAGAGCTTCCCTATTAAATTAAGCTTTGGAGAGATTAATAAAAAA 3325

RESULT 9
AX451929
LOCUS AX451929 3429 bp DNA linear PAT 03-JUL-2002
DEFINITION Sequence 26 from Patent WO0226825.
ACCESSION AX451929
VERSION AX451929.1 GI:21698752
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1
REFERENCE
AUTHORS Baughn, M.R., Graul, R.C., Walia, N.K., Gandhi, A.R., Hafalia, A.J.,
Ramkumar, J., Tribouley, C.M., Thornton, M., Kallick, D.A., Yao, M.G.,
Elliott, V.S., Burford, N., Khan, F.A., Yue, H., Lu, Y., Arvizu, C.,
Roopa, R., Nguyen, D.B., Lee, E.A., Lu, D.A., Isom, C.H., Walsh, R.T. and
Policky, J.L.
TITLE G-protein coupled receptors
JOURNAL Patent: WO 0226825-A 26 04-APR-2002;
Incyte Genomics, Inc. (US)
FEATURES
source Location/Qualifiers
1..3429
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
/note="Incyte ID No: 7479890CB1"

ORIGIN
Query Match 90.1%; Score 2241; DB 6; Length 3429;
Best Local Similarity 99.7%; Pred. No. 0;
Matches 2255; Conservative 0; Mismatches 5; Indels 1; Gaps 1;
QY 133 AGCGCCCTCCACACACACCGCATCTGGGAATTTGAGCTNGACACTTCAGCCAGCTGAG 192
Db 1078 ATCGCCCTCCACACACACCGCATCTGGGAATTTGAGCT -GACACCTTCAGCCAGCTGAG 1136
QY 193 CTCCTCGCAAGCCCTGGATCTTAGCTGGAACCCATCCGGTCCATCCACCCCTGAGGCCTT 252
Db 1137 CTCCTCGCAAGCCCTGGATCTTAGCTGGAACCCATCCGGTCCATCCACCCGAGGCCTT 1196
QY 253 CTCACCCCTGCACTCCCTGGTCAAGCTGGACCTGACAGACACACACAGCTGACACACTGCC 312
Db 1197 CTCACCCCTGCACTCCCTGGTCAAGCTGGACCTGACAGACACACAGCTGACACACTGCC 1256
QY 313 CTGCTGTGACTTGGGGCTTGATGCATCTGAAGCTCAAGGAAACCTTGCTCTCTCCCA 372
Db 1257 CTGCTGTGACTTGGGGCTTGATGCATCTGAAGCTCAAGGAAACCTTGCTCTCTCCCA 1316
QY 373 GGCCTTCTCCAGGACAGTTTCCCAACTGAGGATCTCGGAGGTGCCTTATGCCATCCA 432
Db 1317 GGCCTTCTCCAGGACAGTTTCCCAAACTGAGGATCTCGGAGGTGCCTTATGCCATCCA 1376
QY 493 GTGCTGTCCCTATGGAGTGTGTCAGACTTCTTCAAGGCCTCTGGGCACTGGGAGGCTGA 492
Db 1377 GTGCTGTCCCTATGGAGTGTGTCAGACTTCTTCAAGGCCTCTGGGCACTGGGAGGCTGA 1436
QY 493 AGACCTTCACTTGTATGATGAGGAGTCTTCAAAAAGGCCCTCGGCCCTCTTGGCAGACA 552
Db 1437 AGACCTTCACTTGTATGATGAGGAGTCTTCAAAAAGGCCCTCGGCCCTCTTGGCAGACA 1496
QY 553 AGCAGAACCACTATGACCAGGACCTGGATGAGCTCCAGCTGGAGATGGAGGACTCAAA 612
Db 1497 AGCAGAACCACTATGACCAGGACCTGGATGAGCTCCAGCTGGAGATGGAGGACTCAAA 1556

QY 613 GCCACACCCAGTGTCCAGTGTAGCCCTTACTCCAGGCCCTTCAAGCCCTGTGAGTACCT 672
Db 1557 GCCACACCCAGTGTCCAGTGTAGCCCTTACTCCAGGCCCTTCAAGCCCTGTGAGTACCT 1616
QY 673 CTTTGAAGCTTGGGGCATCCGCTGGCGGTGTGGGCATCGTGTGCTCTCGGTGCTCTG 732
Db 1617 CTTTGAAGCTTGGGGCATCCGCTGGCGGTGTGGGCATCGTGTGCTCTCGGTGCTCTG 1676
QY 733 CAATGGACTGTGTGCTGACCGTGTTCGCTGGGGGGCTGCCCCCTGCCCGCTCAA 792
Db 1677 CAATGGACTGTGTGCTGACCGTGTTCGCTGGGGGGCTGCCCCCTGCCCGCTCAA 1736
QY 793 CTTTGTGTAGGTGCGATTGACGGCGCAACACTTGTGACTGGCATTTCTGTGGCTTCT 852
Db 1737 GTTTGTGTAGGTGCGATTGACGGCGCAACACTTGTGACTGGCATTTCTGTGGCTTCT 1796
QY 853 AGCCTCAGTCGATGCCCTGACCTTTGGTCACTTCTCTGAGTACGGAGCCCTGGAGAC 912
Db 1797 AGCCTCAGTCGATGCCCTGACCTTTGGTCACTTCTCTGAGTACGGAGCCCTGGAGAC 1856
QY 913 GGGGCTAGGCTGCGGGGCACTGGCTTCTCGGCAGTACTTGGTTCGAGGATCTCGGTGCT 972
Db 1857 GGGGCTAGGCTGCGGGGCACTGGCTTCTCGGCAGTACTTGGTTCGAGGATCTCGGTGCT 1916
QY 973 GCTGCTCACTCTGCGCGCAGTGCAGCGTCTCCGCTCTCTGTGTCCGGGCTATGG 1032
Db 1917 GCTGCTCACTCTGCGCGCAGTGCAGCGTCTCCGCTCTCTGTGTCCGGGCTATGG 1976
QY 1033 GAAGTCCCTTCCCTGGGACGGTTCGAGCAGGGGTCTTAGGCTGCCCTGGACACTGGCAG 1092
Db 1977 GAAGTCCCTTCCCTGGGACGGTTCGAGCAGGGGTCTTAGGCTGCCCTGGCAGACTGGCAG 2036
QY 1093 GCTGCGCGCGCACTGCCCTCGGCTCAGTGGGAGAAATACGGGGCTCCCCACTCTGCT 1152
Db 2037 GCTGCGCGCGCACTGCCCTCGGCTCAGTGGGAGAAATACGGGGCTCCCCACTCTGCT 2096
QY 1153 GCCCTACGCGCCACTTGAGGGTTCAGCCAGCAGCCCTGGGCTTTCACCGTGGCCCTGGTAT 1212
Db 2097 GCCCTACGCGCCACTTGAGGGTTCAGCCAGCAGCCCTGGGCTTTCACCGTGGCCCTGGTAT 2156
QY 1213 GATGAATCTCTTCTGTTTCTGTCGTGGCGGGTCCCTACATCAAACTGTACTGTGACCT 1272
Db 2157 GATGAATCTCTTCTGTTTCTGTCGTGGCGGGTCCCTACATCAAACTGTACTGTGACCT 2216
QY 1273 GCGCGGGGCGACTTTGAGGCGCTGTGGCACTGGCCATGCTGAGGACAGCTGGCCCTGGCT 1332
Db 2217 GCGCGGGGCGACTTTGAGGCGCTGTGGCACTGGCCATGCTGAGGACAGCTGGCCCTGGCT 2276
QY 1333 CATCTTCGACAGCGGGCTCTTACTGTCCGCTGGCCCTTCTCAGCTTCGCTCCATGCT 1392
Db 2277 CATCTTCGACAGCGGGCTCTTACTGTCCGCTGGCCCTTCTCAGCTTCGCTCCATGCT 2336
QY 1393 GGGCTCTTCCCTGTACGCCCGAGCGGCTCAAGTCTGTCTGTGCTGTGCTGTGCTGCCCT 1452
Db 2337 GGGCTCTTCCCTGTACGCCCGAGCGGCTCAAGTCTGTCTGTGCTGTGCTGTGCTGCCCT 2396
QY 1453 GCCTGCTGCTCAACCCACTGCTGTACTCTTCAACCCCACTTCCGGGATGACCT 1512
Db 2397 GCCTGCTGCTCAACCCACTGCTGTACTCTTCAACCCCACTTCCGGGATGACCT 2456
QY 1513 TCGGCGGCTTCGGGCCCGCGCAGGGGACTCAGGGGCCCTTAGCCTATGCTGGGGCGGGGA 1572
Db 2457 TCGGCGGCTTCGGGCCCGCGCAGGGGACTCAGGGGCCCTTAGCCTATGCTGGGGCGGGGA 2516
QY 1573 GCTGAGAAAGCTCTCTGTGATTTCTACCCAGGCCCTGTAGCTTCTCTGATGTGGATCT 1632
Db 2517 GCTGAGAAAGCTCTCTGTGATTTCTACCCAGGCCCTGTAGCTTCTCTGATGTGGATCT 2576
QY 1633 CATCTGGAAGCTTCTGAAGCTGGCGGCCCTTGGGCTGGAGACCTATGCTTCCCTC 1692
Db 2577 CATCTGGAAGCTTCTGAAGCTGGCGGCCCTTGGGCTGGAGACCTATGCTTCCCTC 2636
QY 1693 AGTGACCCCTCATCTCTCTGTACAGCAGCCAGGGGGCCCCCAGGGCTGGAGGCGCCTTGTGT 1752

Db	2637	AGTGACCTCATCTCTGTGACAGCCAGGGGCCCCCAGGCTGGAGGCGACCATGTGT	2696
Qy	1753	AGAGCAGAGGGGAACCACTTTGGGAACCCCAACCCCTCCATGGATGAGAACTGCTGCT	1812
Db	2697	AGAGCAGAGGGGAACCACTTTGGGAACCCCAACCCCTCCATGGATGAGAACTGCTGCT	2756
Qy	1813	GAGGCAGAGGGATCTACGCCAGCAGGTGGAGGCTTGTTCAGGGGTGGCGCTTCAGCC	1872
Db	2757	GAGGCAGAGGGATCTACGCCAGCAGGTGGAGGCTTGTTCAGGGGTGGCGCTTCAGCC	2816
Qy	1873	CTCTGGCTTGGCTTTGCTTTCACAGCTGTAATATCCCTCCCCATCTTCTTCTTCCCTC	1932
Db	2817	CTCTGGCTTGGCTTTGCTTTCACAGCTGTAATATCCCTCCCCATCTTCTTCTTCCCTC	2876
Qy	1933	TCTTCCCTTCTCTCTCCCTCGGTGAATGATGGCTGCTTCTTAAACAATAACAACCA	1992
Db	2877	TCTTCCCTTCTCTCTCCCTCGGTGAATGATGGCTGCTTCTTAAACAATAACAACCA	2936
Qy	1993	AAACTCAGCAGTGTGATCTATAGCAGGATGCCAGTACCTGGCTCCACTGATCACCTCT	2052
Db	2937	AAACTCAGCAGTGTGATCTATAGCAGGATGCCAGTACCTGGCTCCACTGATCACCTCT	2996
Qy	2053	CTCTGTGACCATCACCAACGGGCTCTTGGCTGGCTTTCCCTTGGCTTCTTCTCAGC	2112
Db	2997	CTCTGTGACCATCACCAACGGGCTCTTGGCTGGCTTTCCCTTGGCTTCTTCTCAGC	3056
Qy	2113	TTACCTTGATCTGGGCTCTTCTTGTCTATGCTGAAGCTGTGGACGAGACCTTGA	2172
Db	3057	TTACCTTGATCTGGGCTCTTCTTGTCTATGCTGAAGCTGTGGACGAGACCTTGA	3116
Qy	2173	CTTTTGTCTGCTTAAGGGAATGAGGGAATGAGGGAATGAGGGAATGAGGGAATGATC	2232
Db	3117	CTTTTGTCTGCTTAAGGGAATGAGGGAATGAGGGAATGAGGGAATGAGGGAATGATC	3176
Qy	2233	AGGCACAGTGGACAGGGAGACCTCACAGAGAAAGGCTTGAAGGTGATTTCCCGTGTGA	2292
Db	3177	AGGCACAGTGGACAGGGAGACCTCACAGAGAAAGGCTTGAAGGTGATTTCCCGTGTGA	3236
Qy	2293	CTCATGGATAGGATACAAATGTGTTCATGTACATTAATCTTGACATATGCATGCAT	2352
Db	3237	CTCATGGATAGGATACAAATGTGTTCATGTACATTAATCTTGACATATGCATGCAT	3296
Qy	2353	AAAGACTTCTTAAATAAGCTTTGGAAGATTAATAA 2393	
Db	3297	AAAGACTTCTTAAATAAGCTTTGGAAGATTAACA 3337	

RESULT 10	
AX301831	
LOCUS	AX301831
DEFINITION	Sequence 10 from Patent WO0185768.
ACCESSION	AX301831
VERSION	AX301831.1
KEYWORDS	GI:17382891
SOURCE	Homo sapiens (human)
ORGANISM	Homo sapiens
REFERENCE	1. Gu, W.
AUTHORS	Novel g-protein coupled receptors and uses therefor
TITLE	Patent: WO 0185768-A 10 15-NOV-2001;
JOURNAL	Millennium Pharmaceuticals, Inc. (US)
FEATURES	Location/Qualifiers
source	1. .3492
	/organism="Homo sapiens"
	/mol_type="unassigned DNA"
	/db_xref="taxon:9606"
	104. .3007
	/note="unnamed protein product"
	/codon_start=1
	/protein_id="CAD13081.1"
CDS	

/db_xref="GI:17382892"	
/translation="MESPGLRALMLCAALCASRRAGAPQPGPPTACAPACHQED	
GIMLSDCSELGILSAVPGDLPLTAYLDLSNNLTLEQPLGFLHRLFEELRLSGNHL	
SHIPQAFSGYLSIKILMNQNLGGIPAEALWELPSQLSLRLDANLISLVPKSFEG	
LSSRLHLDNDNALTEIPRALNLPALQMTLALNRISHIPDYAFQNLISLVVHLH	
NNRIQHLGTHSFEGLNLETLDNLYNKLOFPVAIRTLGLRLQELGFNNNIKATPEKA	
FMGNLLQTHFYDNPQIOFVGRSAFOYLPKHLTSLNGAMDIOEFPDLKGTTSLEILT	
LTRAGRLPSPGMCQOLPLRVLELSHNOYIEELPSLRHCOKLEIEIGLQHNRIWIGAD	
TFQSLQALDLSWNAIRSIHPEASTLSLVKLDLTQNLTLPLAGLGLMHLKL	
KGNLALSAFQKDSFKPLRIEVPYATCCPYGCASFRRKASGWEADLHLDEESS	
XRLGLLAQENHEDQDELQEMEDSKHPFSVQCSTPGPKPECEYLFESGIRL	
AVWAILVLLVCLNGVLLTVFAGGPAPLPVFKVFGAIGANTLITGSCGLASVDAL	
TFQFSYGARWETGLCRATGFLAVLGSASVLLLTAAVQCSVSVSCVRAYGKSL	
LGSVRAGVLGLALGLAALPLASVEYGCASPLCLPYAPREGOPALGHTVALVMN	
SFLPLVAGAYIKLYCDLPRGDFEAVWDCMWRHVAWLIIFADGLLYCPVAFSLFASML	
GFLPVTPEAVKSVLLVVLPLACNLPLLYLLFNPHFRDDRRLRPGDSDGSPALYAAA	
GELEKSCDSQALVAFSDVLDLLEASGRPFQLETFYFPFVTLISCCQPGAPRLRG	
SHCVEPEGNHFGNPQPSMDGELLRLAEGSTPAGGLSGGGGFGPSPGLAFASHV"	
ORIGIN	
Query Match 50.08; Score 2238; DB 6; Length 3492;	
Best Local Similarity 99.6%; Pred. No. 0;	
Matches 2246; Conservative 7; Mismatches 2; Indels 1; Gaps 1;	
Qy	133 AGCGGCTCCAAACACACACCCGATCTGGGAAATTGGAGCTNGACACCTTCAGCCAGCTGAG 192
Db	1238 ATCGGCTTCCAAACACACACCCGATCTGGGAAATTGGAGCT-GACACCTTCAGCCAGCTGAG 1296
Qy	193 CTCCCTCAAGCCCTTGGATCTTAGCTGGAAACGCAATCCGGTCCATCCACCTTAGGGCTT 252
Db	1297 CTCCCTCAAGCCCTTGGATCTTAGCTGGAAACGCAATCCGGTCCATCCACCTTAGGGCTT 1356
Qy	253 CTCCACCTCGATCCCTCGTCAAGCTGGAGCTGACAGACAAACAGCTGACCACTGCC 312
Db	1357 CTCCACCTCGATCCCTCGTCAAGCTGGAGCTGACAGACAAACAGCTGACCACTGCC 1416
Qy	313 CTGGCTGGACTTGGGGCTTGTATGATCTGAAGCTCAAAAGGAACTTGTCTCTCTCCA 372
Db	1417 CTGGCTGGACTTGGGGCTTGTATGATCTGAAGCTCAAAAGGAACTTGTCTCTCTCCA 1476
Qy	373 GGCTTCTTCAAGGACAGTTTCCAAAACCTGAGGATCTCTGAGGTGCTTATGCTTACCA 432
Db	1477 GGCTTCTTCAAGGACAGTTTCCAAAACCTGAGGATCTCTGAGGTGCTTATGCTTACCA 1536
Qy	433 GTGCTGTCCCTATGGGATGTGTCCAGCTTCTTCAAGGCTCTTGGGAGTGGAGGCTGA 492
Db	1537 GTGCTGTCCCTATGGGATGTGTCCAGCTTCTTCAAGGCTCTTGGGAGTGGAGGCTGA 1596
Qy	493 AGACCTTTCACCTTGATGATGAGGATCTTCAAAAAGGCCCTTGGGCTCTCTTGGCAGACA 552
Db	1597 AGACCTTTCACCTTGATGATGAGGATCTTCAAAAAGGCCCTTGGGCTCTCTTGGCAGACA 1656
Qy	553 AGCAGAGAACCACTATGACCCAGGACCTGGATGAGCTTCCAGCTGGAGATGGAGGACTCAAA 612
Db	1657 AGCAGAGAACCACTATGACCCAGGACCTGGATGAGCTTCCAGCTGGAGATGGAGGACTCAAA 1716
Qy	613 GCACACCCAGTGTCCAGTGTAGCCCTTCTTCAAGGCCCTTCAAGCCCTGTGAGTACCT 672
Db	1717 GCACACCCAGTGTCCAGTGTAGCCCTTCTTCAAGGCCCTTCAAGCCCTGTGAGTACCT 1776
Qy	673 CTTTGAAGCTGGGGATCCGCTGGCGGTGGCCCATCGTGTGTGCTCTCCGTGCTCTG 732
Db	1777 CTTTGAAGCTGGGGATCCGCTGGCGGTGGCCCATCGTGTGTGCTCTCCGTGCTCTG 1836
Qy	733 CAATGGAATGGTGTGTGTGACCCGTGTGTGTGGCGGCTTGGCCCCCTGCCCGGFTCAA 792
Db	1837 CAATGGAATGGTGTGTGTGACCCGTGTGTGTGGCGGCTTGGCCCCCTGCCCGGFTCAA 1896
Qy	793 GTTTGTGGTAGGTGGATTCAGCGCCCAACACCTTGTGAGTGGGATTCCTGTGGCTTCT 852
Db	1897 GTTTGTGGTAGGTGGATTCAGCGCCCAACACCTTGTGAGTGGGATTCCTGTGGCTTCT 1956
Qy	853 AGCTCAGTGCATGCCCTGACCTTTTGTGTCAGTTTCTCTGAGTACGAGGCCCGCTGGGAGAC 912

Db 1957 AGCCTCAGTCGATGCGCCTTACGCTTTGGTCAAGTTCCTGAGTACGAGGCCGCTGGAGAC 2016
Qy 913 GGGGCTAGGCTGCGGGGCACTGGCTTCTGSCAGTACTTGGGTGGAGGCATCGGTGCT 972
Db 2017 GGGGCTAGGCTGCGGGGCACTGGCTTCTGSCAGTACTTGGGTGGAGGCATCGGTGCT 2076
Qy 973 GCTGCTCACTCTGGCGCAGTSCAGTGCAGAGCGTCTCCGTCTCTGTCCTGTCGCGGCCATATGG 1032
Db 2077 GCTGCTCACTCTGGCGCAGTSCAGTGCAGAGCGTCTCCGTCTCTGTCCTGTCGCGGCCATATGG 2136
Qy 1033 GAACTCCCGCTCCCTGGGAGGCTTCGAGCAGGGGTCTTAGGCTCCCTGGCACTGGCAGG 1092
Db 2137 GAACTCCCGCTCCCTGGGAGGCTTCGAGCAGGGGTCTTAGGCTCCCTGGCACTGGCAGG 2196
Qy 1093 GCTGCGCCGCGACCTGCGCCTGCGCTCAGTGGGAGAAACAGGGGCTCCCACTCTGCGCT 1152
Db 2197 GCTGCGCCGCGACCTGCGCCTGCGCTCAGTGGGAGAAACAGGGGCTCCCACTCTGCGCT 2256
Qy 1153 GCCCTACGCGCACCTGAGGGTTCAGCCAGCAGCCCTGGGCTTTCACCGTGGCCCTGGTGAT 1212
Db 2257 GCCCTACGCGCACCTGAGGGTTCAGCCAGCAGCCCTGGGCTTTCACCGTGGCCCTGGTGAT 2316
Qy 1213 GATGAACCTCTTCTGCTTCTGCTGCTGCGCGGCTGCTTACATCAAACTGTGTGACCT 1272
Db 2317 GATGAACCTCTTCTGCTTCTGCTGCTGCGCGGCTGCTTACATCAAACTGTGTGACCT 2376
Qy 1273 GCCGCGGGCGACCTTTCAGGGCGCTGGGACCTGCGCCATGCTGAGGACGCTGGCCTGGCT 1332
Db 2377 GCCGCGGGCGACCTTTCAGGGCGCTGGGACCTGCGCCATGCTGAGGACGCTGGCCTGGCT 2436
Qy 1333 CATCTTCGACAGCGGCTCCTCTACTGTCGCGGCTTCTCTCAGCTTCGCGCTCCATGCT 1392
Db 2437 CATCTTCGACAGCGGCTCCTCTACTGTCGCGGCTTCTCTCAGCTTCGCGCTCCATGCT 2496
Qy 1393 GGGCTCTTCCCTGTCAGCGCGAGGCGCTCAAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1452
Db 2497 GGGCTCTTCCCTGTCAGCGCGAGGCGCTCAAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2556
Qy 1453 GCCTGCGCTCAACCCACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1512
Db 2557 GCCTGCGCTCAACCCACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2616
Qy 1513 TCGCGGCTTCGGCCCGCGCAGGGAAGTCAAGGCGGCTTCTGCTGCTGCTGCTGCTGCTGCTGCT 1572
Db 2617 TCGCGGCTTCGGCCCGCGCAGGGAAGTCAAGGCGGCTTCTGCTGCTGCTGCTGCTGCTGCTGCT 2676
Qy 1573 GCTGAGAGAGCTCTCTGTGATTTACCCAGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1632
Db 2677 GCTGAGAGAGCTCTCTGTGATTTACCCAGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2736
Qy 1633 CATCTGGAAGCTTCTGAGCTGGGCGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1692
Db 2737 CATCTGGAAGCTTCTGAGCTGGGCGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2796
Qy 1693 AGTGACCTCTATCTCTGTCAGCAGCAGGCGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1752
Db 2797 AGTGACCTCTATCTCTGTCAGCAGCAGGCGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2856
Qy 1753 AGAGCCAGAGGGGAACCACTTTGGGAAACCCCAACCTTCCATGGATGAGAACTGCTGCT 1812
Db 2857 AGAGCCAGAGGGGAACCACTTTGGGAAACCCCAACCTTCCATGGATGAGAACTGCTGCT 2916
Qy 1813 GAGGCGAGAGGATCTAGCCAGCAGGTGGAGGCTTGTGAGGGGTGGGCTTTCAGCC 1872
Db 2917 GAGGCGAGAGGATCTAGCCAGCAGGTGGAGGCTTGTGAGGGGTGGGCTTTCAGCC 2976
Qy 1873 CTCTGGCTTGGCTTGTGCTTACAGCTGTAAATATCCCTCCCACTTCTTCTTCTTCTTCTTCTTCT 1932
Db 2977 CTCTGGCTTGGCTTGTGCTTACAGCTGTAAATATCCCTCCCACTTCTTCTTCTTCTTCTTCTTCT 3036
Qy 1933 TCTTCCCTTCTCTCTCCCTCGGTGAATGAGTGTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1992

Db 3037 TCTTCCCTTCTCTCTCCCTCGGTGAATGATGGCTGCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 3096
Qy 1993 AAATCAGCAGTGTGATCTATAGCAGGATGCCCAAGTACCTGGCTCCCACTGATCACTCTCT 2052
Db 3097 AAATCAGCAGTGTGATCTATAGCAGGATGCCCAAGTACCTGGCTCCCACTGATCACTCTCT 3156
Qy 2053 CTCTGTGACCATCAACCAACCGGCTCTTGGGCTGGCTTCTTGGCTTCTTGGCTTCTTGGCTTCTTGGCT 2112
Db 3157 CTCTGTGACCATCAACCAACCGGCTCTTGGGCTGGCTTCTTGGCTTCTTGGCTTCTTGGCTTCTTGGCT 3216
Qy 2113 TTCACTTGTACTTGGGCTCTTCTTGTCTGATCTGATCTGATCTGATCTGATCTGATCTGATCTGATCTGATCT 2172
Db 3217 TTCACTTGTACTTGGGCTCTTCTTGTCTGATCTGATCTGATCTGATCTGATCTGATCTGATCTGATCTGATCT 3276
Qy 2173 CTTTGTCTGCTTAAAGGAAATGAGGAAATGAGGAAATGAGGAAATGAGGAAATGAGGAAATGAGGAAATGAGGAAAT 2232
Db 3277 CTTTGTCTGCTTAAAGGAAATGAGGAAATGAGGAAATGAGGAAATGAGGAAATGAGGAAATGAGGAAATGAGGAAAT 3336
Qy 2233 AGGCGACAGTGGACAGGAGACCTCACAGAGAAAGGCTGGAAGGTGATTTCCCGTGTGA 2292
Db 3337 AGGCGACAGTGGACAGGAGACCTCACAGAGAAAGGCTGGAAGGTGATTTCCCGTGTGA 3396
Qy 2293 CTCTGATAGATACAAAATGCTTCCATGTACCATTAATCTTGACATATGCCATGAT 2352
Db 3397 CTCTGATAGATACAAAATGCTTCCATGTACCATTAATCTTGACATATGCCATGAT 3456
Qy 2353 AAAGACTTCTTAAATAAAGCTTTTGAAGAGATT 2388
Db 3457 AAARACTTCTTAAATAAAGCTTTTGAAGAGATT 3492

RESULT 11
LOCUS CQ870634 3273 bp DNA linear PAT 13-SEP-2004
DEFINITION Sequence 43 from Patent WO2004073657.
ACCESSION CQ870634
VERSION CQ870634.1 GI:52000145
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS Aziz, N., Gish, K.C., Wilson, K.E. and Zlotnik, A.
TITLE Methods of diagnosis of cancer, composition and methods of
screening for modulators of cancer
JOURNAL Patent: WO 2004073657-A 43 02-SEP-2004;
PROTEIN DESIGN LABS, INC. (US)
FEATURES
Location/Qualifiers
source 1..3273
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

ORIGIN
Query Match 89.8%; Score 2231.4; DB 6; Length 3273;
Best Local Similarity 99.7%; Pred. No. 0;
Matches 2246; Conservative 0; Mismatches 6; Indels 1; Gaps 1;
Qy 133 AGCGGCTTCCAAACAACCGCATCTGGGAAATTTGGAGTNGACACCTTTCAGCCAGCTGAG 192
Db 1022 ATCGGCTTCCAAACAACCGCATCTGGGAAATTTGGAGCT-GACACCTTTCAGCCAGCTGAG 1080
Qy 193 CTCCTGCAAGCCCTGGATCTTAGCTGGAAAGCCATCCGGTCCATCCACCTTTCAGGCGCTT 252
Db 1081 CTCCTGCAAGCCCTGGATCTTAGCTGGAAAGCCATCCGGTCCATCCACCTTTCAGGCGCTT 1140
Qy 253 CTCACACCTGCATCTCCCTGGTCAAGCTGGAGCTGACAGAAACCCAGCTGACACACTGCC 312
Db 1141 CTCACACCTGCATCTCCCTGGTCAAGCTGGAGCTGACAGAAACCCAGCTGACACACTGCC 1200
Qy 313 CTTGGCTGGACTTGGGGCTTTGATGCACTCTGAACTCAAGGGAACTTGTCTCTCTCCA 372

Db 1201 CCTGCTGGACCTTGGGGGCTTGATGCATCTGAAGCTCAAAGGAACCTTGCTCTCTCCCA 1260
Qy 373 GGCCTTCTCCAGGACATTTCCAAAACCTGAGGATCCTGGAGTGCCCTTATGCTACCA 432
Db 1261 GGCCTTCTCCAGGACATTTCCAAAACCTGAGGATCCTGGAGTGCCCTTATGCTACCA 1320
Qy 433 GTGCTGTCCCTATGGGATGTGTGCAGCTTCTTCAAGGCCTCTGGGAGTGGGAGGCTGA 492
Db 1321 GTGCTGTCCCTATGGGATGTGTGCAGCTTCTTCAAGGCCTCTGGGAGTGGGAGGCTGA 1380
Qy 493 AGACCTTCACTGTATGATGAGGAGTCTTCAAAAAGGCCCTTGGGCCTCTCTTGGCAGACA 552
Db 1381 AGACCTTCACTGTATGATGAGGAGTCTTCAAAAAGGCCCTTGGGCCTCTCTTGGCAGACA 1440
Qy 553 AGCAGAGAACCATATGACACGAGGACCTGGATGAGCTCCAGCTGGAGATGGAGGACTCAAA 612
Db 1441 AGCAGAGAACCATATGACACGAGGACCTGGATGAGCTCCAGCTGGAGATGGAGGACTCAAA 1500
Qy 613 GCCACACCCAGTGTCCAGTGTAGCCCTACTCCAGGCCCTTCAAGGCCCTGTGAGTACCT 672
Db 1501 GCCACACCCAGTGTCCAGTGTAGCCCTACTCCAGGCCCTTCAAGGCCCTGTGAGTACCT 1560
Qy 673 CTTTGAAGCTGGGCAATCCGCTGGCGTGTGGGCCATCGTGTGCTCTCCGTGCTGTG 732
Db 1561 CTTTGAAGCTGGGCAATCCGCTGGCGTGTGGGCCATCGTGTGCTCTCCGTGCTGTG 1620
Qy 733 CAATGGACTGTGCTGACCGGTGTGCTGGGGGGCTGCGCCCTGCCCCCGGTCAA 792
Db 1621 CAATGGACTGTGCTGACCGGTGTGCTGGGGGGCTGCTGCCCTGCCCCCGGTCAA 1680
Qy 793 GTTTGTGTAGTGTGGATGTGAGGCGGCAACACCTTGACTGGCATTTCTGTGACCTTCT 852
Db 1681 GTTTGTGTAGTGTGGATGTGAGGCGGCAACACCTTGACTGGCATTTCTGTGACCTTCT 1740
Qy 853 AGCCTCAGTGCATGCCCTGACCTTTTGGTCACTTCTTGAGTACGAGGACCCCTGGAGAC 912
Db 1741 AGCCTCAGTGCATGCCCTGACCTTTTGGTCACTTCTTGAGTACGAGGACCCCTGGAGAC 1800
Qy 913 GGGGCTAGGCTGCCGGGCACTGGCTTCTGGCAGTACTTGGTTCGAGGACCTGGTGCT 972
Db 1801 GGGGCTAGGCTGCCGGGCACTGGCTTCTGGCAGTACTTGGGTCGAGGACCTGGTGCT 1860
Qy 973 GCTGCTCACTCTGGCCGAGTGCAGTGCAGCGTCTCCGTCTCTGTGTCCGGGCTATGG 1032
Db 1861 GCTGCTCACTCTGGCCGAGTGCAGTGCAGGCTCTCCGTCTCTGTGTCCGGGCTATGG 1920
Qy 1033 GAACTCCCTCTCTGGGCACTGTCAGCAGGGGTCTTAGGCTGCCCTGGCACTGGCAGG 1092
Db 1921 GAACTCCCTCTCTGGGCACTGTCAGCAGGGGTCTTAGGCTGCCCTGGCACTGGCAGG 1980
Qy 1093 GCTGGCCGCCCACTGGCCCTGGCTCAGTGGGAGAAATACGGGGCTCCCACTCTGCCT 1152
Db 1981 GCTGGCCGCCCGCTGGCCCTGGCTCAGTGGGAGAAATACGGGGCTCCCACTCTGCCT 2040
Qy 1153 GCCCTAGCGCCACTGAGGGTCAGCAGCAGCCCTGGGCTTACCGTGGCCCTGGTGAT 1212
Db 2041 GCCCTAGCGCCACTGAGGGTCAGCAGCAGCCCTGGGCTTACCGTGGCCCTGGTGAT 2100
Qy 1213 GATGAACCTCTCTGTTTCTGGTCTGGCCGGTGCCTTACATCAAACTGACTGTGACCT 1272
Db 2101 GATGAACCTCTCTGTTTCTGGTCTGGCCGGTGCCTTACATCAAACTGACTGTGACCT 2160
Qy 1273 GCCCGGGGGCACTTTGAGGCCGTGTGGGACTGGCCCATGTGAGGCACGTGGCCCTGGCT 1332
Db 2161 GCCCGGGGGCACTTTGAGGCCGTGTGGGACTGGCCCATGTGAGGCACGTGGCCCTGGCT 2220
Qy 1333 CATCTTCGACAGCGGCTCTCTACTGTTCGGTGGCTTCTCTCAGCTTCGGCTCCATGCT 1392
Db 2221 CATCTTCGACAGCGGCTCTCTACTGTTCGGTGGCTTCTCTCAGCTTCGGCTCCATGCT 2280
Qy 1393 GGGCCTCTTCCCTGACGCCCGAGGCCGTCAAGTCTGTCTCTGCTGGTGTGCTGCCCT 1452
Db 2281 GGGCCTCTTCCCTGACGCCCGAGGCCGTCAAGTCTGTCTCTGCTGTGCTGTGCTGCCCT 2340

Qy 1453 GCCTGCTGCTCAACCCACCTGCTGTACCTGCTCTTCAACCCCCACCTTCCGGGATGACCT 1512
Db 2341 GCCTGCTGCTCAACCCACCTGCTGTACCTGCTCTTCAACCCCCACCTTCCGGGATGACCT 2400
Qy 1513 TCGCGGCTTTCGGCCCGCCGCGCAGGGGACTCAGGCGCCCTTAGCCTATGCTGGGCGCGGGA 1572
Db 2401 TCGCGGCTTTCGGCCCGCCGCGCAGGGGACTCAGGCGCCCTTAGCCTATGCTGGGCGCGGGA 2460
Qy 1573 GCTGGAGAAGAGCTCTCTGTGATTTCTA CCGAGGCCCTCGGTAGCCTTCTCTGATGTGATCT 1632
Db 2461 GCTGGAGAAGAGCTCTCTGTGATTTCTA CCGAGGCCCTCGGTAGCCTTCTCTGATGTGATCT 2520
Qy 1633 CATTTCTGGAAGCTTCTCAAGCTGGCGGCCCTTGGGCTGGAGACCTATGCTTCCCTC 1692
Db 2521 CATTTCTGGAAGCTTCTCAAGCTGGCGGCCCTTGGGCTGGAGACCTATGCTTCCCTC 2580
Qy 1693 AGTGACCTCTCATCTCTCTGTGTCAGCAGCCAGGGGCCCTCCAGGCTGGAGGCGACCAATTGTGT 1752
Db 2581 AGTGACCTCTCATCTCTCTGTGTCAGCAGCCAGGGGCCCTCCAGGCTGGAGGCGACCAATTGTGT 2640
Qy 1753 AGAGCCAGAGGGGAAACACCTTTTGGGAAACCCCCAAACCTCTCATGTGATGGAGAACTGTGCT 1812
Db 2641 AGAGCCAGAGGGGAAACACCTTTTGGGAAACCCCCAAACCTCTCATGTGATGGAGAACTGTGCT 2700
Qy 1813 GAGGCGAGAGGGATCTACGCCAGCAGGTGGAGGCTTGT CAGGGGGTGGCGCTTTCAGCC 1872
Db 2701 GAGGCGAGAGGGATCTACGCCAGCAGGTGGAGGCTTGT CAGGGGGTGGCGCTTTCAGCC 2760
Qy 1873 CTCTGGCTTGGCCCTTTGCTTTCAACAGTGTAAATATCCCTCCCAATTTCTCTTCCCTC 1932
Db 2761 CTCTGGCTTGGCCCTTTGCTTTCAACAGTGTAAATATCCCTCCCAATTTCTCTTCCCTC 2820
Qy 1933 TCTTCCCTTCTCTCTCTCCCTCGGTGAATGATGTGCTTCTTAAACAAATAACAACCA 1992
Db 2821 TCTTCCCTTCTCTCTCTCCCTCGGTGAATGATGTGCTTCTTAAACAAATAACAACCA 2880
Qy 1993 AAACCTCAGCAGTGTGATCTATAGCAGATGCCAGTACCTGGCTCCCACTGATCACCTCT 2052
Db 2881 AAACCTCAGCAGTGTGATCTATAGCAGATGCCAGTACCTGGCTCCCACTGATCACCTCT 2940
Qy 2053 CTCTGTGACCATCACCAACGGGTGCTCTTGGGCTGGCTTTCCCTTGGGCTTCCCTCAGC 2112
Db 2941 CTCTGTGACCATCACCAACGGGTGCTCTTGGCTGGCTTTCCCTTGGGCTTCCCTCAGC 3000
Qy 2113 TTCACTTTGATACCTGGGCTTCTCTTGTCACTGTCTGAAGCTGTGGACGAGACCTTGA 2172
Db 3001 TTCACTTTGATACCTGGGCTTCTCTTGTCACTGTCTGAAGCTGTGGACGAGACCTTGA 3060
Qy 2173 CTTTGTGCTCTTAAAGGAAATGAGGGAAGTAAAGACAGTGAAGGGGTGGAGGGTGTGATC 2232
Db 3061 CTTTGTGCTCTTAAAGGAAATGAGGGAAGTAAAGACAGTGAAGGGGTGGAGGGTGTGATC 3120
Qy 2233 AGGCGACAGTGGACAGGGAGACCTCACAGAAAGGCTCGAAAGGTGATTTCCCGTGA 2292
Db 3121 AGGCGACAGTGGACAGGGAGACCTCACAGAAAGGCTCGAAAGGTGATTTCCCGTGA 3180
Qy 2293 CTCAATGATAGGATACAAAATGTGTTCATGTACATTAATCTTGAATATGCAATATGCCATGCAT 2352
Db 3181 CTCAATGATAGGATACAAAATGTGTTCATGTACATTAATCTTGAATATGCAATATGCCATGCAT 3240
Qy 2353 AAAGACTTCTCTATTAATAAAGCTTTTGGAGAG 2385
Db 3241 AAAGACTTCTCTATTAATAAAGCTTTTGGAGAG 3273

RESULT 12
E4291.7
LOCUS
DEFINITION Novel gonadotropin receptor.
ACCESSION E4291.7
VERSION E4291.1 GI:18627553
KEYWORDS JP 2000125884 -A/5.

E4291.7 3041 bp DNA linear PAT 31-JAN-2002

SOURCE	unidentified	Db	1257	CTTTGAAAGCTGGGGCATCCGCTGGCGGTGTGGGCCATCGTGTGCTCTCCGTGCTCTG	1316
ORGANISM	unclassified.	Qy	733	CAATGGACTGTGTGCTGACCGTGTTCGCTGGCGGGCTCCGCCCTCCCGCTCCCGGTCAA	792
REFERENCE	1 (bases 1 to 3041)	Db	1317	CAATGGACTGTGTGCTGACCGTGTTCGCTGGCGGGCTCCGCCCTCCCGGTCAA	1376
AUTHORS	Heikopu, Y. C. and Supekku, P. J. V. D.	Qy	793	GTTTGTGGTAGTCCGATTGTCAGCGCCAAACACTTGTGATGGCATTTCTGTGGCTTCT	852
TITLE	Novel gonadotropin receptor	Db	1377	GTTTGTGGTAGTCCGATTGTCAGCGCCAAACACTTGTGATGGCATTTCTGTGGCTTCT	1436
JOURNAL	Patent: JP 2000125884-A 5 09-MAY-2000;	Qy	853	AGCTCAGTCATGCCCTGACCTTTGGTTCAGTTCCTGAGTACGGAGCCCGCTGGGAGAC	912
COMMENT	AK20 NOBEL NV	Db	1437	AGCTCAGTCATGCCCTGACCTTTGGTTCAGTTCCTGAGTACGGAGCCCGCTGGGAGAC	1496
	OS Unidentified	Qy	913	GGGGTAGGCTGCCGGGGCACTGGCTTCTCGCAGTACTTGGGTCCGAGGACATCGGTGCT	972
	PN JP 2000125884-A/5	Db	1497	GGGGTAGGCTGCCGGGGCACTGGCTTCTCGCAGTACTTGGGTCCGAGGACATCGGTGCT	1556
	PD 05-MAY-2000	Qy	973	GCTGCTCACTCTGGCCGAGTCAGTGCAGGTTCCTGCTCTCTGCTGTCGGGCTATGG	1032
	PR 05-FEB-1999 JP 1999029120	Db	1557	GCTGCTCACTCTGGCCGAGTCAGTGCAGGTTCCTGCTCTCTGCTGTCGGGCTATGG	1616
	PR 06-FEB-1998 EP 98200357.6, 27-JUL-1998 EP 98202519.9, PR	Qy	1033	GAAGTCCCTCTCCCTGGGACGCTTCGAGCAGGGGTCTTAGGCTGCTGAGCACTGGCAGG	1092
	24-SEP-1998 EP 98203213.8	Db	1617	GAAGTCCCTCTCCCTGGGACGCTTCGAGCAGGGGTCTTAGGCTGCTGAGCACTGGCAGG	1676
	PI YUDEISU CHRISTINA HEIKOPU, PETORIYUSU JOHANNES VAN DEL SUPEKKU	Qy	1093	GCTGGCGCGCACTGCCCTCGGCTCAGTGGGAGAAATACGGGGCTCCCCACTCTGCCT	1152
	PC C12N15/09, A61K38/00, A61P15/08, A61P43/00, C07K14/72, C12N1/15, PC	Db	1677	GCTGGCGCGCACTGCCCTCGGCTCAGTGGGAGAAATACGGGGCTCCCCACTCTGCCT	1736
	C12N1/19,	Qy	1153	GCCCTAGCGCCACCTCAGGGTTCAGCAGCAGCCCTGGGCTTTCACCGTGGCCCTGGTGAT	1212
	PC C12N1/21, C12N5/10, C12P21/02, C12Q1/68, G01N33/15, G01N33/50, PC	Db	1737	GCCCTAGCGCCACCTCAGGGTTCAGCAGCAGCCCTGGGCTTTCACCGTGGCCCTGGTGAT	1796
	G01N33/566,	Qy	1213	GATGAACTCCTTCTGTTCTGTTCTGTTCTGTTCTGTTCTGTTCTGTTCTGTTCTGTTCT	1272
	PC C12N15/00, A61K37/02, C12N5/00	Db	1797	GATGAACTCCTTCTGTTCTGTTCTGTTCTGTTCTGTTCTGTTCTGTTCTGTTCTGTTCT	1856
	CC Strandedness: Single;	Qy	1273	CCCGCGGGCGACCTTTGAGGCGCTGTGGGACTGTGGCATGTGTGAGGACAGTGGCGCTGCT	1332
	CC Topology: Linear;	Db	1857	CCCGCGGGCGACCTTTGAGGCGCTGTGGGACTGTGGCATGTGTGAGGACAGTGGCGCTGCT	1916
	FT Key	Qy	1333	CATCTTCGACAGCGGGCTCTTACTGTCCGCTGGCCCTTCTCAGCTTCGCTCCATGCT	1392
	FT source	Db	1917	CATCTTCGACAGCGGGCTCTTACTGTCCGCTGGCCCTTCTCAGCTTCGCTCCATGCT	1976
FEATURES	Location/Qualifiers	Qy	1393	GGGCTCTTCTCCTGTACGCCCGAGCGGTCAAGTCTGTCTGCTGTGTGTGTGTGTGTGTGT	1452
source	1..3041	Db	1977	GGGCTCTTCTCCTGTACGCCCGAGCGGTCAAGTCTGTCTGCTGTGTGTGTGTGTGTGTGT	2036
	/organism='unidentified'	Qy	1453	GCCTGCTGCTCAACCCACTGCTGCTCTTCAACCCCACTTCCGGGATGACCT	1512
	/no_type='genomic DNA'	Db	2037	GCCTGCTGCTCAACCCACTGCTGCTCTTCAACCCCACTTCCGGGATGACCT	2096
	/db_xref='caxon:32644'	Qy	1513	TCGGCGGCTTCGGCCCGCGCAGGGGACTCAGGGCCCTAGCCTATGCTGGCGCGGGGA	1572
ORIGIN		Db	2097	TCGGCGGCTTCGGCCCGCGCAGGGGACTCAGGGCCCTAGCCTATGCTGGCGCGGGGA	2156
	Query Match 89.3%; Score 2219.6; DB 6; Length 3041;	Qy	1573	GCTGAGAGAGCTCCTGTGATTCTACCCAGGCCCTGTGTAGCTTCTCTGATGTGGATCT	1632
	Best Local Similarity 99.7%; Pred No. 0;	Db	2157	GCTGAGAGAGCTCCTGTGATTCTACCCAGGCCCTGTGTAGCTTCTCTGATGTGGATCT	2216
	Matches 2255; Conservative 0; Mismatches 4; Indels 3; Gaps 3;	Qy	1633	CATTCTGGAAGCTTCTGAAGCTGGGCGGCCCTGGGCTGGAGACCTATGCTTCCCTC	1692
Qy	133 AGCGGCTCCCAACACACCGCATCTGGGAAATGTGAGCTNGACACCTTCAGCCAGCTGAG	Db	2217	CATTCTGGAAGCTTCTGAAGCTGGGCGGCCCTGGGCTGGAGACCTATGCTTCCCTC	2276
Db	718 ATCGGCTCCCAACACACCGCATCTGGGAAATGTGAGCT-GACACCTTCAGCCAGCTGAG	Qy	1693	AGTGACCTCATCTCCTGTGACAGCAGGGGCGGCCCAAGGCTGGAGGCGACGCAATTGTGT	1752
Qy	193 CTCCTGCAAGCCCTGGATCTTAGTGGNAGCCATCCGGTCCATCCACCTCAGGCGCTT	Db	2277	AGTGACCTCATCTCCTGTGACAGCAGGGGCGGCCCAAGGCTGGAGGCGACGCAATTGTGT	2336
Db	777 CTCCTGCAAGCCCTGGATCTTAGTGGNAGCCATCCGGTCCATCCACCTCAGGCGCTT	Qy	1753	AGAGCCAGAGGGGAAACCACTTTGGGAAACCCCAACCTTCATGATGAGAACTGCTGCT	1812
Qy	253 CTCACCTTGCACTCCCTGGTCAAGCTGGACCTGACAGCAACACAGCTGACACACTGCC				
Db	837 CTCACCTTGCACTCCCTGGTCAAGCTGGACCTGACAGCAACACAGCTGACACACTGCC				
Qy	313 CTTGCTGCACTTGGGGCTTGATGATCTGAAGCTCAAAAGGAACTTGTCTCTCCCA				
Db	897 CTTGCTGCACTTGGGGCTTGATGATCTGAAGCTCAAAAGGAACTTGTCTCTCCCA				
Qy	373 GGCCTTCTCAAGGACAGTTTCCCAAACTGAGGATCTGAGGTCCTTATGCTTACCA				
Db	957 GGCCTTCTCAAGGACAGTTTCCCAAACTGAGGATCTGAGGTCCTTATGCTTACCA				
Qy	433 GTGCTGCTCCATGAGTGTGTCAGCTTCTTCAAGGCTCTGGGAGTGGGAGGCTGA				
Db	1017 GTGCTGCTCCATGAGTGTGTCAGCTTCTTCAAGGCTCTGGGAGTGGGAGGCTGA				
Qy	493 AGACCTTCACTTGATGAGGAGTCTTCAAAAAGGCCCTTGGGCTCTCTTGGCAGACA				
Db	1077 AGACCTTCACTTGATGAGGAGTCTTCAAAAAGGCCCTTGGGCTCTCTTGGCAGACA				
Qy	553 AGCAGAGAACCACTATGACAGGACCTGATGAGCTCAGCTGGAGATGGAGACTCAAA				
Db	1137 AGCAGAGAACCACTATGACAGGACCTGATGAGCTCAGCTGGAGATGGAGACTCAAA				
Qy	613 GCCACACCCAGTGTCCAGTGTAGCCCTTCTCCAGGCCCTTCAAGCCCTGTGAGTACT				
Db	1197 GCCACACCCAGTGTCCAGTGTAGCCCTTCTCCAGGCCCTTCAAGCCCTGTGAGTACT				
Qy	673 CTTTGAAGCTGGGGCATCCGCTGGCGGTGTGGGCCATCGTGTGCTCTCCGTGCTCTG				

QY	1273	GGCGGGGGCGACTTTGAGGCGCGTGTGGGACTGCGCCATGTTGAGGACGTTGGCCTGGCT	1332
Db	1857	GGCGGGGGCGACTTTGAGGCGCGTGTGGGACTGCGCCATGTTGAGGACGTTGGCCTGGCT	1916
QY	1333	CATCTTCGACAGCGGCTCTACTGTCGCCGTGGCCTTCCTCAGCTTCGCCTCCATGCT	1392
Db	1917	CATCTTCGACAGCGGCTCTACTGTCGCCGTGGCCTTCCTCAGCTTCGCCTCCATGCT	1976
QY	1393	GGGCTCTTCCCTGTGACGCCGAGGCGCTCAAGTCTGTCTCTGCTGGTGTCTGCCCT	1452
Db	1977	GGGCTCTTCCCTGTGACGCCGAGGCGCTCAAGTCTGTCTCTGCTGGTGTCTGCCCT	2036
QY	1453	GCTGCTGCTCAACCCACTGCTGTACTGCTCTTCAACCCCACTTCCGGGATGACT	1512
Db	2037	GCTGCTGCTCAACCCACTGCTGTACTGCTCTTCAACCCCACTTCCGGGATGACT	2096
QY	1513	TCGGCGGCTTCGGCCCGCGCAGGAGACTCAGGGCCCTAGCCTATGCTCGCGCGGGGA	1572
Db	2097	TCGGCGGCTTCGGCCCGCGCAGGAGACTCAGGGCCCTAGCCTATGCTCGCGCGGGGA	2156
QY	1573	GCTGAGAGAGCTCTGTGATTCTACCCAGGCCCTGGTAGCTTCTCTGATGTGATCT	1632
Db	2157	GCTGAGAGAGCTCTGTGATTCTACCCAGGCCCTGGTAGCTTCTCTGATGTGATCT	2216
QY	1633	CATTCTGGAAGCTTCTGAAGCTGGGCGGCCCTGGGTGGAGACTATGGCTTCCCTC	1692
Db	2217	CATTCTGGAAGCTTCTGAAGCTGGGCGGCCCTGGGTGGAGACTATGGCTTCCCTC	2276
QY	1693	AGTGACCTCATCTCTGTGACAGCAGGAGGCCGCCAGGCTGGAGGCGCCATTGTGT	1752
Db	2277	AGTGACCTCATCTCTGTGACAGCAGGAGGCCGCCAGGCTGGAGGCGCCATTGTGT	2336
QY	1753	AGAGCCAGAGGGGAACCACTTTGGGAACCCCAACCTCCATGGATGAGAACTCTGCT	1812
Db	2337	AGAGCCAGAGGGGAACCACTTTGGGAACCCCAACCTCCATGGATGAGAACTCTGCT	2396
QY	1813	GAGGCGAGAGGATCTAGCCAGCAGGTGGAGGCTTGTACGGGGTGGCGCTTTAGCC	1872
Db	2397	GAGGCGAGAGGATCTAGCCAGCAGGTGGAGGCTTGTACGGGGTGGCGCTTTAGCC	2455
QY	1873	CTCTGGCTTGGCTTGTCTACACGTGTAAATATCCCTCCCATCTCTCTCCCTC	1932
Db	2456	CTCTGGCTTGGCTTGTCTACACGTGTAAATATCCCTCCCATCTCTCTCCCTC	2515
QY	1933	TCTTCCCTTCTCTCTCCCTCGGTGAATGATGGCTTCTTAAACAATAACAACA	1992
Db	2516	TCTTCCCTTCTCTCTCCCTCGGTGAATGATGGCTTCTTAAACAATAACAACA	2575
QY	1993	AACTCAGCAGTGTATCTATAGCAGGATGGCCAGTACCTGGCTCCACTGATCACTCT	2052
Db	2576	AACTCAGCAGTGTATCTATAGCAGGATGGCCAGTACCTGGCTCCACTGATCACTCT	2635
QY	2053	CTCCTGTACCATCACCAACGGGTG-CCTCTTGGCTGGCTTTCCTTGGCCCTCCTCAG	2111
Db	2636	CTCCTGTACCATCACCAACGGGTG-CCTCTTGGCTGGCTTTCCTTGGCCCTCCTCAG	2695
QY	2112	CTTCACTTGTATCTGGGCTCTTCTCTGTCTATGCTGAAGCTGTGGACACGACCTGG	2171
Db	2696	CTTCACTTGTATCTGGGCTCTTCTCTGTCTATGCTGAAGCTGTGGACACGACCTGG	2755
QY	2172	ACTTTTGTCTTAAAGGAAATAGGGAAGTAAAGACAGTGAAGGGGTGGAGGTTGAT	2231
Db	2756	ACTTTTGTCTTAAAGGAAATAGGGAAGTAAAGACAGTGAAGGGGTGGAGGTTGAT	2815
QY	2232	CAGGCGACAGTGGACAGGAGACCTCACAGAGAAAGGCTGGAAGTGTATTCCTGCTG	2291
Db	2816	CAGGCGACAGTGGACAGGAGACCTCACAGAGAAAGGCTGGAAGTGTATTCCTGCTG	2875
QY	2292	ACTCATGATAGGATACAAAATGTGTCCATGTACCATTAATCTTTGACATATGCCATGA	2351
Db	2876	ACTCATGATAGGATACAAAATGTGTCCATGTACCATTAATCTTTGACATATGCCATGA	2935
QY	2352	TAAAGACTTCTTATTAATAAGCTTTTGGAGAGATTAATAA	2393

Db	2936	TAAAGACTTCTTATTAATAAGCTTTTGGAGAGATTACACA	2977
RESULT 14			
AX527920			
LOCUS	AX527920	3381 bp	DNA linear PAT 21-NOV-2002
DEFINITION	Sequence 5 from Patent WO0229038.		
ACCESSION	AX527920		
VERSION	AX527920.1	GI:25172361	
KEYWORDS			
SOURCE	Homo sapiens (human)		
ORGANISM	Homo sapiens		
REFERENCE			
AUTHORS	Herrmann, J.L., Rastelli, L. and Shimkets, R.A.		
TITLE	Novel proteins and nucleic acids encoding same and antibodies directed against these proteins		
JOURNAL	Patent: WO 0229038-A 5 11-APR-2002;		
	Curagen Corporation (US)		
FEATURES	Location/Qualifiers		
source	1..3381		
	/organism="Homo sapiens"		
	/mol_type="unassigned DNA"		
	/db_xref="taxon:9606"		
ORIGIN			
Query Match	89.3%;	Score 2219.6;	DB 6; Length 3381;
Best Local Similarity	99.7%;	Pred. No. 0;	
Matches 2255;	Conservative 0;	Mismatches 4;	Indels 3; Gaps 3;
QY	133	AGCGCCCTCCAAACACAAACCGCATCTGGGAAATTTGGAGCTNGACACCTTTCAGCCAGCTCAG	192
Db	1058	ATCGCCCTCCAAACACAAACCGCATCTGGGAAATTTGGAGCT-GACACCTTCAGCCAGCTCAG	1116
QY	193	CTCCCTGCAAGCCCTGGATCTTAGCTGGAACGCCATCCGGTCCATCCACCCCGAGGCTT	252
Db	1117	CTCCCTGCAAGCCCTGGATCTTAGCTGGAAGGCCATCCGGTCCATCCACCCCGAGGCTT	1176
QY	253	CTCCACCTGACATCCCTGGTCAAGCTGGAGCTGACAGAACCCAGCTGACCACTGCC	312
Db	1177	CTCCACCTGACATCCCTGGTCAAGCTGGAGCTGACAGAACCCAGCTGACCACTGCC	1236
QY	313	CCTGGCTGATCTGGGGCTTTGATCATCTGAAGCTCAAAGGGAACCTTGCTCTCTCCA	372
Db	1237	CCTGGCTGATCTGGGGCTTTGATCATCTGAAGCTCAAAGGGAACCTTGCTCTCTCCA	1296
QY	373	GGCCTTCTCCAAGGACAGTTCCTCCAAAACCTGAGGATCCTGGAGTGCCTTATGCTACCA	432
Db	1297	GGCCTTCTCCAAGGACAGTTCCTCCAAAACCTGAGGATCCTGGAGTGCCTTATGCTACCA	1356
QY	433	GTGCTGCTCCTATGGGATGTGTCCAGCTTCTTCAAGSCCTCTGGGAGTGGGAGGCTGA	492
Db	1357	GTGCTGCTCCTATGGGATGTGTCCAGCTTCTTCAAGSCCTCTGGGAGTGGGAGGCTGA	1416
QY	493	AGACCTTCACCTTGATGATGAGGAGTCTTCAAAAAGGCGCTTGGGCTCTTGCAGACA	552
Db	1417	AGACCTTCACCTTGATGATGAGGAGTCTTCAAAAAGGCGCTTGGGCTCTTGCAGACA	1476
QY	553	AGCAGAGAACCACTATGACACGAGCCTGGATGAGCTCCAGCTGGAGATGGAGGACTCAAA	612
Db	1477	AGCAGAGAACCACTATGACACGAGCCTGGATGAGCTCCAGCTGGAGATGGAGGACTCAAA	1536
QY	613	GCCACACCCAGTGTCCAGTGTAGCCCTACTTCCAGGCGCCCTTCAAGCCCTGTGAGTACT	672
Db	1537	GCCACACCCAGTGTCCAGTGTAGCCCTACTTCCAGGCGCCCTTCAAGCCCTGTGAGTACT	1596
QY	673	CTTTGAAAGCTGGGGCATCCGCTGGCGTGTGGGCCATCGTGTGCTCTCCGTGCTCTG	732
Db	1597	CTTTGAAAGCTGGGGCATCCGCTGGCGTGTGGGCCATCGTGTGCTCTCCGTGCTCTG	1656
QY	733	CAATGGACTGTGCTGTGCTGACCGTGTTCGCTGGGGGCGCTGCCCGCCCTGCCCGGTCAA	792

Db 1657 CAATGGACTGGTGTCTGACCGTGTTCGCTGGCGGGCTGCCCCCTGCCCCCGGTCAA 1716
 QY GTTTGTGTAGGTGGATTCAGCGGCCAACACACTTGACTGGCATTTCTGTGGCTTCT 852
 Db 1717 GTTTGTGTAGGTGGATTCAGCGGCCAACACACTTGACTGGCATTTCTGTGGCTTCT 1776
 QY AGCCTCAGTCAGATGCCCTGACCTTTGGTTCAGTTCCTGAGTACGGAGCCCGCTGGGAGAC 912
 Db 1777 AGCCTCAGTCAGATGCCCTGACCTTTGGTTCAGTTCCTGAGTACGGAGCCCGCTGGGAGAC 1836
 QY GGGGTAGGCTGCGGGCCACTGGCTTCTG3CAGTACTTGGGTG3GAGGATCGGTGCT 972
 Db 1837 GGGGTAGGCTGCGGGCCACTGGCTTCTG3CAGTACTTGGGTG3GAGGATCGGTGCT 1896
 QY GCTGCTACTCTGGCGCAGTCAGTCAGCGTCTCCGTCTCTGTGTGTCGGGCCATATGG 1032
 Db 1897 GCTGCTACTCTGGCGCAGTCAGTCAGCGTCTCCGTCTCTGTGTGTCGGGCCATATGG 1956
 QY GAAATCCCCCTCCCTGGGCAGCGTTTCGAGCAGGGGTCTTAGGCTGSCCTGGCACTGGCAGG 1092
 Db 1957 GAAATCCCCCTCCCTGGGCAGCGTTTCGAGCAGGGGTCTTAGGCTGSCCTGGCACTGGCAGG 2016
 QY GCTGSCCGCAGCTGCCCTGGCTCAGTGGGAGAAATACGGGGCTTCCCACTCTGCCT 1152
 Db 2017 GCTGSCCGCAGCTGCCCTGGCTCAGTGGGAGAAATACGGGGCTTCCCACTCTGCCT 2076
 QY GCCCTACGGGCACCTGAGGGTCAGCCAGCAGCCCTGGGCTTACCGTGGGCCCTGGTGAT 1212
 Db 2077 GCCCTACGGGCACCTGAGGGTCAGCCAGCAGCCCTGGGCTTACCGTGGGCCCTGGTGAT 2136
 QY GATGAACCTCTTCTGTTTCTGTTTCTGTCGTCGTCGCGGTCTACAACTGTACTGTGACCT 1272
 Db 2137 GATGAACCTCTTCTGTTTCTGTTTCTGTCGTCGTCGCGGTCTACAACTGTACTGTGACCT 2196
 QY GCCCGGGGCGACTTTGAGGCGGTGTGGGACTGCGGCCATGTGTGAGGCACTGTGGCTGGCT 1332
 Db 2197 GCCCGGGGCGACTTTGAGGCGGTGTGGGACTGCGGCCATGTGTGAGGCACTGTGGCTGGCT 2256
 QY CATCTTCGACAGCGGGCTCTCTACTGTCCCGTGGCCCTTCTCAGCTTCGGCTCGATGCT 1392
 Db 2257 CATCTTCGACAGCGGGCTCTCTACTGTCCCGTGGCCCTTCTCAGCTTCGGCTCGATGCT 2316
 QY GGGCTCTTCCCTGTACGCGCCGAGGCGGTCAAGTCTGTCTGTCTGTGGTGTGCTGCCCT 1452
 Db 2317 GGGCTCTTCCCTGTACGCGCCGAGGCGGTCAAGTCTGTCTGTCTGTGGTGTGCTGCCCT 2376
 QY GCCTGCTGCTCAACCCCACTGCTGTACTGCTCTTCAACCCCACTTTCGGGATGACCT 1512
 Db 2377 GCCTGCTGCTCAACCCCACTGCTGTACTGCTCTTCAACCCCACTTTCGGGATGACCT 2436
 QY TCGGGGGTTCGGGCCCGCCGACGGGACTCAGGGCCCTAGCCCTATGCTGGCGCGGGGA 1572
 Db 2437 TCGGGGGTTCGGGCCCGCCGACGGGACTCAGGGCCCTAGCCCTATGCTGGCGCGGGGA 2496
 QY GCTGAGAGAGCTCTCTGTATTTACCCAGCCCTGTAGCTTCTCTGATGTGGATCT 1632
 Db 2497 GCTGAGAGAGCTCTCTGTATTTACCCAGCCCTGTAGCTTCTCTGATGTGGATCT 2556
 QY CATCTGGAAGCTTCTGAACTGAGCTGGCGGCCCTTGGCTGGAGACCTATGGCTTCCCTC 1692
 Db 2557 CATCTGGAAGCTTCTGAACTGAGCTGGCGGCCCTTGGCTGGAGACCTATGGCTTCCCTC 2616
 QY AGTGACCTCATCTCTGTTCAGCAGCAGCGGGGCCCGCCAGGCTGGAGGACGCCATTTGT 1752
 Db 2617 AGTGACCTCATCTCTGTTCAGCAGCAGCGGGGCCCGCCAGGCTGGAGGACGCCATTTGT 2676
 QY AGAGCCAGAGGGGAACCACTTTTGGGAACCCCAACCTTCCATGTATGAGAACTGCTGCT 1812
 Db 2677 AGAGCCAGAGGGGAACCACTTTTGGGAACCCCAACCTTCCATGTATGAGAACTGCTGCT 2736
 QY GAGGCGAGAGGGATCTAGCCAGCAGGTGGAGGCTTGTTCAGGGGTGGCGGCTTTCAGCC 1872

Db 2737 GAGGCGAGAGGATCTTAGCCAGCAGGTGGAGGCTTGTTCAGGGGTGGC- GCTTTTCAGCC 2795
 QY CTCTGGCTTGGCTTTCCTTTCACACGTTGTAATATCCTCCCTCCCATTTCTTCTTCCCTC 1932
 Db 2796 CTCTGGCTTGGCTTTCCTTTCACACGTTGTAATATCCTCCCTCCCATTTCTTCTTCCCTC 2855
 QY TCTTCCCTTTCCTCTCTCCCTCGTGAATGATGGCTTCTTAAACAATAACAACA 1992
 Db 2856 TCTTCCCTTTCCTCTCTCCCTCGTGAATGATGGCTTCTTAAACAATAACAACA 2915
 QY AAATCAGCAGTGTGATCTATAGCAGGATGCGCCAGTACCTGGCTCCATGATCACCTCT 2052
 Db 2916 AAATCAGCAGTGTGATCTATAGCAGGATGCGCCAGTACCTGGCTCCATGATCACCTCT 2975
 QY CTCTGTGACCATCACCAACGGGTG- CCTCTTGGCTTGGCTTTCCTTGGCTTTCCTCAG 2111
 Db 2976 CTCTGTGACCATCACCAACGGGTGCTTCTTGGCTTGGCTTTCCTTGGCTTTCCTCAG 3035
 QY CTTTCACTTGTACTTGGGCTTCTTCTTGTCTGTGATGCTGTGAAGCTGTGGACAGACCTGG 2171
 Db 3036 CTTTCACTTGTACTTGGGCTTCTTCTTGTCTGTGATGCTGTGAAGCTGTGGACAGACCTGG 3095
 QY ACTTTTGTCTTAAAGGAAATGAGGAAAGTAAAGACAGTGAAGGGGTGGAGGTTGAT 2231
 Db 3096 ACTTTTGTCTTAAAGGAAATGAGGAAAGTAAAGACAGTGAAGGGGTGGAGGTTGAT 3155
 QY CAGGCGACGTGGACGGGAGACCTCAGAGAAAGCCCTGGAAGGTGATTTCCCGGTG 2291
 Db 3156 CAGGCGACGTGGACGGGAGACCTCAGAGAAAGCCCTGGAAGGTGATTTCCCGGTG 3215
 QY ACTCATGATAGGATCAAAATGTGTTCCATGTACCAATTAATCTTGACATATGCCATGCA 2351
 Db 3216 ACTCATGATAGGATCAAAATGTGTTCCATGTACCAATTAATCTTGACATATGCCATGCA 3275
 QY TAAAGACTTCTTAAATAAAGCTTTGGAAGAGATTACACA 2393
 Db 3276 TAAAGACTTCTTAAATAAAGCTTTGGAAGAGATTACACA 3317

RESULT 15
 Locus AX549295 3119 bp DNA linear PAT 26-NOV-2002
 DEFINITION Sequence 580 from Patent WO02061087.
 ACCESSION AX549295
 VERSION AX549295.1 GI:25813958
 KEYWORDS
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE
 1 Burmer, G.C., Roush, C.L. and Brown, J.P.
 Antigenic peptides, such as for G protein-coupled receptors
 (GPCRs), antibodies thereto, and systems for identifying such
 antigenic peptides
 Patent: WO 02061087-A 580 08-AUG-2002;
 Lifespan Biosciences, Inc. (US)
 FEATURES
 source 1. 3119
 /organism="Homo sapiens"
 /mol_type="unassigned DNA"
 /db_xref="taxon:9606"

ORIGIN
 Query Match 83.9%; Score 2085.8; DB 6; Length 3119;
 Best local Similarity 99.9%; Pred. No. 0;
 Matches 2096; Conservative 0; Mismatches 2; Indels 1; Gaps 1;
 QY 133 AGCGGCTCCAAACACACACCGCATCTGGAAATTTGGAGCTNGACACCTTCAGCCAGCTGAG 192
 Db 1020 ATCGGCTCCAAACACACACCGCATCTGGAAATTTGGAGCT-GACACCTTCAGCCAGCTGAG 1078
 QY 193 CTCCCTGCAAGCCCTGGATCTTAGCTGGAACGCGCATCCGGTCCATCCACCCCTGAGGCTT 252

Db 1079 CCCCCGAGGCGCTGGATCTTTAGCTGGAAACCCATCCGGTCCATCCACCTGAGGCGCTT 1138
Qy 253 CTTCCACCCCTGCACTCCCTGGTCAAGCTGAGCTGACAGACAAACAGCTGACACACACTGGC 312
Db 1139 CTTCCACCCCTGCACTCCCTGGTCAAGCTGAGCTGACAGACAAACAGCTGACACACACTGGC 1198
Qy 313 CTTGGCTGCACTTTGGGGCTTGATGATCTGAAGCTCAAGGGAACCTTTGCTCTCTCCCA 372
Db 1199 CTTGGCTGCACTTTGGGGCTTGATGATCTGAAGCTCAAGGGAACCTTTGCTCTCTCCCA 1258
Qy 373 GGCCTTCTCCAGGACAGTTTCCCAAACTGAGGATCTTGAAGGCTCTGGAGGTCCTTATGCCCTACA 432
Db 1259 GGCCTTCTCCAGGACAGTTTCCCAAACTGAGGATCTTGAAGGCTCTGGAGGTCCTTATGCCCTACA 1318
Qy 433 GTGCTGTCTCCATGGGATGTGGCAGCTTCTTCAAGGCTCTGGCAGTGGAGGCTGA 492
Db 1319 GTGCTGTCTCCATGGGATGTGGCAGCTTCTTCAAGGCTCTGGCAGTGGAGGCTGA 1378
Qy 493 AGACCTTCACTTGATGATGAGGAGTCTTCAAAAAGGCCCCCTGGGCTCTCTTGGCAGACA 552
Db 1379 AGACCTTCACTTGATGATGAGGAGTCTTCAAAAAGGCCCCCTGGGCTCTCTTGGCAGACA 1438
Qy 553 AGCAGAGAACCACTATGACAGGACCTGGATGAGCTCCAGCTGGAGATGGAGGACTCAAA 612
Db 1439 AGCAGAGAACCACTATGACAGGACCTGGATGAGCTCCAGCTGGAGATGGAGGACTCAAA 1498
Qy 613 GCCACACCCAGTGTCCAGTGTAGCCCTACTCCAGGCCCTTCAAGCCCTGTGAGTACCT 672
Db 1499 GCCACACCCAGTGTCCAGTGTAGCCCTACTCCAGGCCCTTCAAGCCCTGTGAGTACCT 1558
Qy 673 CTTTGAAGCTGGGCGATCCGCTGGCGGTGGGCCATCGTGTGCTCTCGTGTCTGTG 732
Db 1559 CTTTGAAGCTGGGCGATCCGCTGGCGGTGGGCCATCGTGTGCTCTCGTGTCTGTG 1618
Qy 733 CAATGGACTGTGCTGACCGTGTGCTGGCGGCGCTGCCCTGCCCTGCCCTGCCCTGCA 792
Db 1619 CAATGGACTGTGCTGACCGTGTGCTGGCGGCGCTGCCCTGCCCTGCCCTGCCCTGCA 1678
Qy 793 GTTGTGGTGTGGTGTGGATGGAGGCGCAACACCTTGTAGCTGGCATTTCTGTGGCTTCT 852
Db 1679 GTTGTGTGAGTGTGGATGGAGGCGCAACACCTTGTAGCTGGCATTTCTGTGGCTTCT 1738
Qy 853 AGCCTCAGTGTGCTGACCTTTGGTCACTTCTGAGTACGGAGCGCCGCTGGGAGAC 912
Db 1739 AGCCTCAGTGTGCTGACCTTTGGTCACTTCTGAGTACGGAGCGCCGCTGGGAGAC 1798
Qy 913 GGGGCTAGGCTGCCGGGCGACCTGGCTTCTGGCAGTACTTGGGTCGGAGGATCGGTGCT 972
Db 1799 GGGGCTAGGCTGCCGGGCGACCTGGCTTCTGGCAGTACTTGGGTCGGAGGATCGGTGCT 1858
Qy 973 GCTGCTCACTCTGGCGGAGTGCAGTGCAGCGCTCTCCGTCTCCTGTGTCGGGCGCTATGG 1032
Db 1859 GCTGCTCACTCTGGCGGAGTGCAGTGCAGCGCTCTCCGTCTCCTGTGTCGGGCGCTATGG 1918
Qy 1033 GAAATCCCCCTCCTCTGGGCGAGGTTTCGAGCAGGGGTCTTAGGCTGCTGGCACTGGCAGG 1092
Db 1919 GAAATCCCCCTCCTCTGGGCGAGGTTTCGAGCAGGGGTCTTAGGCTGCTGGCACTGGCAGG 1978
Qy 1093 GCTGCGCGCGCACTGCCCTCGCTCAGTGGGAGAAATACGGGGCTCTCCCACTCTGCT 1152
Db 1979 GCTGCGCGCGCACTGCCCTCGCTCAGTGGGAGAAATACGGGGCTCTCCCACTCTGCT 2038
Qy 1153 GGCCTACGGCGCACCTGAGGTCAGCCAGCAGCCCTGGGCTTTCAGGTGGCGCTGGTGTAT 1212
Db 2039 GGCCTACGGCGCACCTGAGGTCAGCCAGCAGCCCTGGGCTTTCAGGTGGCGCTGGTGTAT 2098
Qy 1213 GATGAATCTCTGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1272
Db 2099 GATGAATCTCTGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2158
Qy 1273 GCCCGGGCGGCACTTTGAGGCGCTGTGGGCACTGCGCCATGGTGGGCACTGCGGCTGGCT 1332

Db 2159 GCCCGGGCGGCACTTTGAGGCGCTGTGGGCACTGCGCCATGTGTGAGGCACTGCGGCTGGCT 2218
Qy 1333 CATCTTCGACAGAGGGCTCTCTACTGTCCGCTGGCTTCTCTCAGCTTTCGCTCCATGCT 1392
Db 2219 CATCTTCGACAGAGGGCTCTCTACTGTCCGCTGGCTTCTCTCAGCTTTCGCTCCATGCT 2278
Qy 1393 GGGCTCTTCTCTGTGACGCGCGAGGCGCTCAAGTCTGTCTCTGTGTGTGTGTGTGTGTGTGT 1452
Db 2279 GGGCTCTTCTCTGTGACGCGCGAGGCGCTCAAGTCTGTCTCTGTGTGTGTGTGTGTGTGTGT 2338
Qy 1453 GCCTGCTGCTCAACCCACCTGCTGTACTGCTTCAACCCCACTTCCGGGATGACCT 1512
Db 2339 GCCTGCTGCTCAACCCACCTGCTGTACTGCTTCAACCCCACTTCCGGGATGACCT 2398
Qy 1513 TCGGCGGCTTCGGGCGCGCGAGGGACTCAGGGGCGCTAGCTATGCTGCGGCGGGGA 1572
Db 2399 TCGGCGGCTTCGGGCGCGCGAGGGACTCAGGGGCGCTAGCTATGCTGCGGCGGGGA 2458
Qy 1573 GCTGGAGAAGAGCTCTCTGTGATTTTACCCAGGCGCTGTGTGTGTGTGTGTGTGTGTGTGT 1632
Db 2459 GCTGGAGAAGAGCTCTCTGTGATTTTACCCAGGCGCTGTGTGTGTGTGTGTGTGTGTGTGT 2518
Qy 1633 CATCTGGAAGCTTCTGAACTGGGCGCGCGCTGGGCTGGAGACTATGGCTTCCCTC 1692
Db 2519 CATCTGGAAGCTTCTGAACTGGGCGCGCGCTGGGCTGGAGACTATGGCTTCCCTC 2578
Qy 1693 AGTGACCTCATCTCTGTGACAGCAGGCGCGCGCTGGGCTGGAGGCGAGCATTTGTGT 1752
Db 2579 AGTGACCTCATCTCTGTGACAGCAGGCGCGCGCTGGGCTGGAGGCGAGCATTTGTGT 2638
Qy 1753 AGAGCCAGAGGGGAACCACTTTGGGAACCCCAACCTTCCATGGATGGAGAACTGCTGCT 1812
Db 2639 AGAGCCAGAGGGGAACCACTTTGGGAACCCCAACCTTCCATGGATGGAGAACTGCTGCT 2698
Qy 1813 GAGGCGAGAGGATCTAGCCAGCAGGTGGAGGCTTGTGAGGGGTGGGGCTTTCAAGC 1872
Db 2699 GAGGCGAGAGGATCTAGCCAGCAGGTGGAGGCTTGTGAGGGGTGGGGCTTTCAAGC 2758
Qy 1873 CTCTGGCTTGGCTTGTGCTTCAACGCTGATAATATCCCTCCCTCATTTCTTCTTCCCTC 1932
Db 2759 CTCTGGCTTGGCTTGTGCTTCAACGCTGATAATATCCCTCCCTCATTTCTTCTTCCCTC 2818
Qy 1933 TCTTCCCTTCTCTCTCCCTCGTGAATGATGGCTTCTTAAAAAATAAATAAATAAATAA 1992
Db 2819 TCTTCCCTTCTCTCTCCCTCGTGAATGATGGCTTCTTAAAAAATAAATAAATAAATAA 2878
Qy 1993 AAACTCAGCAGTGTGATCTATAGCAGGATGGCCAGTACCTGGCTCCACTGATCACCTCT 2052
Db 2879 AAACTCAGCAGTGTGATCTATAGCAGGATGGCCAGTACCTGGCTCCACTGATCACCTCT 2938
Qy 2053 CTCTGTGACCATCACCAACGGGTGCTTGTGGCTGGCTTTCCTTGGCTTCCCTCAGC 2112
Db 2939 CTCTGTGACCATCACCAACGGGTGCTTGTGGCTGGCTTTCCTTGGCTTCCCTCAGC 2998
Qy 2113 TTTACCTTGTACTTGGGCTCTTCTTGTGATGTCTGAAGCTGTGGACAGAGACCTTGA 2172
Db 2999 TTTACCTTGTACTTGGGCTCTTCTTGTGATGTCTGAAGCTGTGGACAGAGACCTTGA 3058
Qy 2173 CTTTTGTCTGCTTAAAGGAAATGAGGGAAGTAAAGACAGTGAAGGGGTGGAGGTTGATC 2232
Db 3059 CTTTTGTCTGCTTAAAGGAAATGAGGGAAGTAAAGACAGTGAAGGGGTGGAGGTTGATC 3118
Qy 2233 A 2233
Db 3119 A 3119

Search completed: April 14, 2005, 15:59:12
Job time : 7120.6 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 14, 2005, 10:59:27 ; Search time 877.045 Seconds
(without alignments)
16779.616 Million cell updates/sec

Title: US-10-664-667-4
Perfect score: 2486
Sequence: 1 taatcagctactatagg.....agtgcacctaattcaatt 2486

Scoring table: IDENTITY_NUC
Gapop 10.0, Gapext 1.0

Searched: 4390206 seqs, 2959870667 residues

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_16Dec04:*
1: Geneseqn1980s:*
2: Geneseqn1990s:*
3: Geneseqn2000s:*
4: Geneseqn2001as:*
5: Geneseqn2001bs:*
6: Geneseqn2002as:*
7: Geneseqn2002bs:*
8: Geneseqn2003as:*
9: Geneseqn2003bs:*
10: Geneseqn2003cs:*
11: Geneseqn2003ds:*
12: Geneseqn2004as:*
13: Geneseqn2004bs:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2485	100.0	2486	6	Aa167922 Human LGR
2	2485	100.0	2486	11	Adn02242 Human par
3	2485	100.0	2486	12	Adk19407 cDNA enco
4	2309.8	92.9	3451	12	Adq76775 Human LGR
5	2309.4	92.9	2322	5	Aaf98722 Human lat
6	2257.4	90.8	2711	6	Aa167924 Partial c
7	2257.4	90.8	2711	11	Adn02245 Human par
8	2257.4	90.8	2711	12	Adk19410 Partial c
9	2255.8	90.7	2988	6	Abz11841 Human pol
10	2255.8	90.7	2988	12	Adm4359 Novel hum
11	2247.8	90.4	3306	12	Adq76773 Human LGR
12	2247.8	90.4	3349	6	Adh50807 Human G-p
13	2247.2	90.4	3325	10	Adh89048 Encoding
14	2243	90.2	2964	10	Adi21981 Novel hum
15	2241	90.1	3429	6	Aad37674 Human G-p
16	2238	90.0	3443	10	Adh89090 Encoding
17	2238	90.0	3492	6	Aa167926 Human LGR
18	2238	90.0	3492	11	Adn02248 Human ful
19	2238	90.0	3492	12	Adk19413 Full leng
20	2238	90.0	3492	12	Adq15073 Human can

21	2231.4	89.8	3273	11	Adn39801	Adn39801 Cancer/an
22	2231.4	89.8	3273	13	Adr46630	Adr46630 Cancer-as
23	2219.6	89.3	3041	2	Aaz40461	Aaz40461 Human gon
24	2219.6	89.3	3381	10	Adg42572	Adg42572 Novel hum
25	2218	89.2	3381	6	Abk49424	Abk49424 DNA enco
26	2123.8	85.4	2168	10	Adi21492	Adi21492 Novel hum
27	2085.8	83.9	3119	8	Abz42549	Abz42549 Human G p
28	1965.8	79.1	3000	6	Aas98067	Aas98067 Human DNA
29	1898	76.3	1899	6	Aa167923	Aa167923 Human LGR
30	1898	76.3	1899	11	Adn02244	Adn02244 Human par
31	1898	76.3	1899	12	Adk19409	Adk19409 ORF of CD
32	1844.2	74.2	1987	2	Aaz40460	Aaz40460 Human gon
33	1754.4	70.6	2208	6	Aa167925	Aa167925 Human LGR
34	1754.4	70.6	2208	11	Adn02247	Adn02247 Human par
35	1754.4	70.6	2208	12	Adk19412	Adk19412 ORF of pa
36	1754.4	70.6	2901	6	Aa167927	Aa167927 Human LGR
37	1754.4	70.6	2901	10	Adc16692	Adc16692 cDNA enco
38	1754.4	70.6	2901	11	Adn02250	Adn02250 Human ful
39	1754.4	70.6	2901	12	Adk19415	Adk19415 ORF of fu
40	1754.4	70.6	3042	10	Adc16698	Adc16698 cDNA enco
41	1754.4	70.6	3438	10	Abt31938	Abt31938 Human bre
42	1752.8	70.5	2786	10	Adc16710	Adc16710 Human G-p
43	1752.8	70.5	3042	10	Adc16700	Adc16700 cDNA enco
44	1752.6	70.5	2487	6	Adh50804	Adh50804 Human G-p
45	1749.6	70.4	2901	10	Adc16695	Adc16695 cDNA with

ALIGNMENTS

RESULT 1
Aa167922
ID Aa167922 standard; cDNA; 2486 BP.
AC Aa167922;
XX
XX
DT 13-MAR-2002 (first entry)
XX
DE Human LGR6 polypeptide encoding cDNA (clone fahr).
XX
KW Large G-protein coupled receptor 6; LGR6; GPCR; immunomodulator;
KW anti-HIV; cytostatic; neuroprotective; antiparkinsonian;
KW anticonvulsant; neuroleptic; antidepressant; tranquilizer; antimigraine;
KW osteopathic; antiarteriosclerotic; vasotropic; hypotensive; cardiant;
KW antiarrhythmic; anorectic; gene therapy; human; ss.
XX
OS Homo sapiens.
XX
FH Key
CDS Location/Qualifiers
FT 2..1903
FT /*tag= a
FT /product= "LGR6 polypeptide"
FT /trans= except= {pos: 170..172, aa: Xaa}
FT /note= "Xaa = unknown"

WO200185768-A2.
15-NOV-2001.
08-MAY-2001; 2001WO-US015002.
08-MAY-2000; 2000US-00566588.
(MILL-) MILLENNIUM PHARM INC.
Gu W;
WPI; 2002-0555584/07.
P-P8DB; AAG66139.

Novel isolated large G-protein coupled receptor 6 polypeptide, useful for diagnosing and treating weight disorder, metabolic disorder, central nervous system disorder, endocrine disorder and cardiovascular disorder.

XX	Example 1; Fig 4; 198pp; English.
PS	The invention provides isolated large G-protein coupled receptor 6 (LGR6)
XX	polypeptides and polynucleotides encoding them. The LGR6 polypeptides and
CC	polynucleotides are useful as targets for developing modulating agents
CC	that regulate a variety of cellular processes, e.g. neural and endocrine
CC	processes, as well as thermogenesis. They are useful for developing novel
CC	diagnostic and therapeutic agents for LGR6 associated disorders such as
CC	weight disorders (anorexia, obesity), eating disorders, acquired
CC	immunodeficiency syndrome (AIDS)-related wasting, cancer-related wasting,
CC	metabolic disorders, central nervous system disorders (e.g. Alzheimer's
CC	disease, Parkinson's disease, Huntington's disease, Gilles de la
CC	Tourette's syndrome, multiple sclerosis, amyotrophic lateral sclerosis,
CC	epilepsy, Jakob-Creutzfeldt disease, depression, schizoprenia,
CC	obsessive-compulsive disorder, anxiety, panic disorder, migraine),
CC	endocrine disorders (hypo and hyperthyroidism, dwarfism, gigantism,
CC	acromegaly), and cardiovascular disorders (atherosclerosis, restenosis,
CC	hypertension, myocardial infarction, arrhythmia). The LGR6 polypeptides,
CC	polynucleotides and antibodies are useful in screening assays, detection
CC	assays (e.g., chromosomal mapping, tissue typing, forensic biology),
CC	predictive medicine (e.g. diagnostic assays, prognostic assays,
CC	monitoring clinical trials and pharmacogenomics), and in methods of
CC	treatment (e.g. therapeutic and prophylactic). The present sequence
CC	represents a human LGR6 polypeptide encoding cDNA
XX	
SQ	Sequence 2486 BP; 455 A; 773 C; 692 G; 565 T; 0 U; 1 Other;
	Query Match 100.0%; Score 2485; DB 6; Length 2486;
	Best Local Similarity 100.0%; Pred. No. 0;
	Matches 2486; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY	1 TAATAGACTCACTATAGGAAAGCTGGTACGCCGTGCAGGTACCGTCCGGAATTCCGG 60
DB	1 TAATAGACTCACTATAGGAAAGCTGGTACGCCGTGCAGGTACCGTCCGGAATTCCGG 60
QY	61 GTCCACCACCGCTCCGTGGAGCGGACGACGAGGTCTGAGCCTGCCGGCTCATCAGCCTC 120
DB	61 GTCCACCACCGCTCCGTGGAGCGGAGCCAGGCTCTGAGCCTGCCGGCTCATCAGCCTC 120
QY	121 TCTTGTCGCCCTAGGGCCTCAAACAACCGCATCTGGAAATTGGAGCTNGACACTT 180
DB	121 TCTTGTCGCCCTAGGGCCTCAAACAACCGCATCTGGGAAATTGGAGCTNMGACACTT 180
QY	181 CAGCAGCTGAGCTCCCTCAAGCCCTTGATCTTAGCTGGAAACGCATCCGGTCCATCCA 240
DB	181 CAGCAGCTGAGCTCCCTGCAAGCCCTTGATCTTAGCTGGAAACGCATCCGGTCCATCCA 240
QY	241 CCCTGAGGCCTTTCTCCACCTGCACCTCCCTGGTCAAGCTGGACCTGACAGAACCAAGCT 300
DB	241 CCCTGAGGCCTTTCTCCACCTGCACCTCCCTGGTCAAGCTGGACCTGACAGAACCAAGCT 300
QY	301 GACCACATGCCCTCGCTGGACTTTGGGGGCTTGATGATCTGAAGCTCAAAGGGAACCT 360
DB	301 GACCACATGCCCTCGCTGGACTTTGGGGGCTTGATGATCTGAAGCTCAAAGGGAACCT 360
QY	361 TGCTCTCTCCAGGCCTTCTCCAGGACAGATTTCCCAAAAATGAGGATCTCTGGAGGTGCC 420
DB	361 TGCTCTCTCCAGGCCTTCTCCAGGACAGATTTCCCAAAAATGAGGATCTCTGGAGGTGCC 420
QY	421 TTATGCTTACCAGTGTGCTCCATATGGATGTGTGCCAGCTTCTTCAAGGCCTCTGGGCA 480
DB	421 TTATGCTTACCAGTGTGCTCCATATGGATGTGTGCCAGCTTCTTCAAGGCCTCTGGGCA 480
QY	481 GTGGGAGGCTGAAGACCTTTCACCTTGATGATGAGGAGTCTTCAAAAAAGGCCCTTGGGCT 540
DB	481 GTGGGAGGCTGAAGACCTTTCACCTTGATGATGAGGAGTCTTCAAAAAAGGCCCTTGGGCT 540
QY	541 CCTTGCACACAGCAGAGAACCACTATGACACGAGCCTGATGAGCTCCAGCTGGAGAT 600
DB	541 CCTTGCACACAGCAGAGAACCACTATGACACGAGCCTGATGAGCTCCAGCTGGAGAT 600
QY	601 GGAGGACTCAAAAGCCACACCCCAGTGTCCAGTGTAGCCCTACTCCAGGCCCTTTCAAGCC 660

Db 2341 TATGCCATGCAATAAGACTTCTATTAATAAAGCTTTGGAAGAGATTAAAAA 2400
Qy 2401 AAAGGGCGCGCTCTAGAGGATCAAGCTTACGTACGCGTGATGCGACGTCATAGCTC 2460
Db 2401 AAAGGGCGCGCTCTAGAGGATCAAGCTTACGTACGCGTGATGCGACGTCATAGCTC 2460
Qy 2461 TTCTATAGTGTCACCTAAATCAATT 2486
Db 2461 TTCTATAGTGTCACCTAAATCAATT 2486
RESULT 3
ID ADK19407
XX ADK19407 standard; cDNA: 2486 BP.
XX
XX AC ADK19407;
XX
XX DT 17-JUN-2004 (first entry)
XX
XX DE CDNA encoding human LGR6 polypeptide #1.
XX
XX KW G-protein coupled receptor; GPCR; large G-protein coupled receptor 6;
KW LGR6; weight disorder; neural disorder; endocrine disorder;
KW cardiovascular disorder; anorectic; anabolic; immunomodulator;
KW neuroprotective; nootropic; antiparkinsonian; anti-HIV; anticonvulsant;
KW antithyroid; osteopathic; cardiant; vasotropic; hypotensive;
KW antiarteriosclerotic; antiarrhythmic; human; gene; ss.
XX
XX OS Homo sapiens.
XX
XX FH Key Location/Qualifiers
XX CDS 2..1903
XX FT /*tag= a
XX FT /partial
XX FT /product= "LGR6 polypeptide #1"
XX FT /note= "This sequence lacks a start codon"
XX FT /transl_except= (pos:170..172,aa:Xaa)
XX FT /note= "Xaa= Unknown"
XX
XX PN US2004058377-A1.
XX
XX PD 25-MAR-2004.
XX
XX PF 18-SEP-2003; 2003US-00664667.
XX
XX PR 06-MAY-1999; 99US-0132896P.
XX PR 08-MAY-2000; 2000US-00566588.
XX PR 08-MAY-2001; 2001US-00851595.
XX
XX PA (MILL-) MILLENNIUM PHARM INC.
XX
XX PI Gu W;
XX
XX DR WPI; 2004-268789/25.
XX DR P-PSDB; ADK19408.
XX
XX PT New large G-protein coupled receptor 6 nucleic acid molecules and
XX polypeptides, useful for diagnosing, preventing or treating diseases
XX associated with aberrant nucleic acid or protein activity, e.g. obesity,
XX anorexia or cachexia.
XX
XX PS Claim 2; SEQ ID NO 4; 145pp; English.
XX
XX CC The present invention relates to the isolation of novel members of the G-
XX protein coupled receptor (GPCR) family designated as large G-protein
XX coupled receptor 6 (LGR6), and the polynucleotide sequences encoding
XX them. The invention also discloses LGR6 fusion proteins, antigenic
XX peptides, anti-LGR6 antibodies, recombinant expression vectors, host
XX cells, and non-human transgenic animals in which an LGR6 gene has been
XX introduced or disrupted. The LGR6 polynucleotide sequences, polypeptide
XX sequences, and antibodies are useful for diagnosing, treating or
XX preventing diseases associated with aberrant LGR6 expression or activity,

CC such as weight disorders (e.g. obesity, anorexia nervosa or cachexia),
CC neural disorders (e.g. central nervous system disorder, dementia,
CC Alzheimer's disease, Parkinson's disease, AIDS-related dementia, multiple
CC sclerosis and Huntington's disease), endocrine disorders (e.g.
CC hypothyroidism, hyperthyroidism or acromegaly), or cardiovascular
CC disorders (e.g. ischaemia-reperfusion injury, coronary artery disease,
CC myocardial infarction, arrhythmia, atherosclerosis, hypertension or
CC congestive heart failure). The LGR6 polynucleotide and polypeptide
CC sequences are also useful as targets for developing modulating agents
CC that regulate a variety of cellular processes, e.g. neural and endocrine
CC processes and thermogenesis. They can be used in screening assays (e.g.
CC chromosome mapping, tissue typing or in forensic biology), or in
CC predictive medicine (e.g. diagnostic assays, prognostic assays,
CC monitoring clinical trials or pharmacogenomics). The present sequence
CC encodes a human LGR6 polypeptide.
XX
XX SQ Sequence 2486 BP; 455 A; 773 C; 692 G; 565 T; 0 U; 1 Other;
Query Match 100.0%; Score 2485; DB 12; Length 2486;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2486; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 TAATACGACTCATTATAGGGAAGCTGGTACGCTGACAGGTACCGTCCGGATTCGCCGG 60
Db 1 TAATACGACTCATTATAGGGAAGCTGGTACGCTGACAGGTACCGTCCGGATTCGCCGG 60
Qy 61 GTCCACCCACCGCTCCGTGGAGCGGACGCGGTCTGAGCCTGCGGCTCATCAGACCTC 120
Db 61 GTCCACCCACCGCTCCGTGGAGCGGACGCGGTCTGAGCCTGCGGCTCATCAGACCTC 120
Qy 121 TCTTGCTGCCCTAGCGGCTCCAAACACACCGCATCTGGGAAATTTGGAGCTNGACACCTT 180
Db 121 TCTTGCTGCCCTAGCGGCTCCAAACACACCGCATCTGGGAAATTTGGAGCTNGACACCTT 180
Qy 181 CAGCAGCTGAGCTCCCTGCAAGCCCTGGATCTTAGCTGGAACGCCATCCGGTCCATCCA 240
Db 181 CAGCAGCTGAGCTCCCTGCAAGCCCTGGATCTTAGCTGGAACGCCATCCGGTCCATCCA 240
Qy 241 CCCTGAGGCTTCTCCACCCCTGCACTCCCTGGTCAAGCTGGACCTGACAGACAACACGCT 300
Db 241 CCCTGAGGCTTCTCCACCCCTGCACTCCCTGGTCAAGCTGGACCTGACAGACAACACGCT 300
Qy 301 GACCACACTGCCCCCTGGCTGGACTTTGGGGGCTTGATGCATCTGAAGCTCAAAGGGAACCT 360
Db 301 GACCACACTGCCCCCTGGCTGGACTTTGGGGGCTTGATGCATCTGAAGCTCAAAGGGAACCT 360
Qy 361 TGCTCTCTCCAGGCTTCTCCAGGACAGTTTCCCAAACTGAGATCTCTGGAGTGCC 420
Db 361 TGCTCTCTCCAGGCTTCTCCAGGACAGTTTCCCAAACTGAGATCTCTGGAGTGCC 420
Qy 421 TTATGCCCTACCAAGTGTCTCCCTATGGATGTGCGCAGCTTCTTCAAGGCTCTGGGCA 480
Db 421 TTATGCCCTACCAAGTGTCTCCCTATGGATGTGCGCAGCTTCTTCAAGGCTCTGGGCA 480
Qy 481 GTGGAGGCTGAAGACCTTCACTTGTATGATGAGGAGTCTTCAAAAAGGCCCTTGGGCT 540
Db 481 GTGGAGGCTGAAGACCTTCACTTGTATGATGAGGAGTCTTCAAAAAGGCCCTTGGGCT 540
Qy 541 CCTTGCCAGACAAGCAGAGAACCACTATGACAGGACCTTGATGAGCTCCAGCTGGAGAT 600
Db 541 CCTTGCCAGACAAGCAGAGAACCACTATGACAGGACCTTGATGAGCTCCAGCTGGAGAT 600
Qy 601 GGAGGACTCAAAGCCACACCCAGTGTCCAGTGTAGCCCTACTCCAGGCCCTTCAAGCC 660
Db 601 GGAGGACTCAAAGCCACACCCAGTGTCCAGTGTAGCCCTACTCCAGGCCCTTCAAGCC 660
Qy 661 CTGTGAGTACCTCTTTGAAAGCTGGGCGCATCCGCTTGCGCTGTGGGCGCATCGTGTGCT 720
Db 661 CTGTGAGTACCTCTTTGAAAGCTGGGCGCATCCGCTTGCGCTGTGGGCGCATCGTGTGCT 720
Qy 721 CTCCGCTGCTGCAATGAGACTGTGGTGTGCTGACCGTGTGCTGGGGGCGCTGCCCCCT 780
Db 721 CTCCGCTGCTGCAATGAGACTGTGGTGTGCTGACCGTGTGCTGGGGGCGCTGCCCCCT 780

Qy	781	GC	CC	CC	CG	GT	CA	AG	TT	TG	TG	TAG	GT	GC	GA	TT	GC	AG	GC	CC	CA	AC	TT	TG	AT	GC	CA	TT	T	C	840		
Db	781	GC	CC	CC	CG	GT	CA	AG	TT	TG	TG	TAG	GT	GC	GA	TT	GC	AG	GC	CC	CA	AC	TT	TG	AT	GC	CA	TT	T	C	840		
Qy	841	CT	GT	GG	CC	TT	TA	GC	CT	CA	GT	CG	AT	GC	CG	CT	GA	CT	TT	TG	GT	CA	GT	T	CT	T	GA	GT	AC	GA	CG	900	
Db	841	CT	GT	GG	CC	TT	TA	GC	CT	CA	GT	CG	AT	GC	CG	CT	GA	CT	TT	TG	GT	CA	GT	T	CT	T	GA	GT	AC	GA	CG	900	
Qy	901	CC	GC	TT	GG	AG	AG	CG	GG	CT	TAG	CT	GC	CG	GC	CA	CT	TG	GG	CT	T	CT	CG	CA	GT	AC	TT	TG	GG	CT	CG	960	
Db	901	CC	GC	TT	GG	AG	AG	CG	GG	CT	TAG	CT	GC	CG	GC	CA	CT	TG	GG	CT	T	CT	CG	CA	GT	AC	TT	TG	GG	CT	CG	960	
Qy	961	GG	CA	T	CG	GT	GT	CT	GT	CT	CA	CT	CT	TG	GG	CG	CA	GT	GC	AG	T	GC	AG	CG	T	CT	CG	T	CT	CG	T	1020	
Db	961	GG	CA	T	CG	GT	GT	CT	GT	CT	CA	CT	CT	TG	GG	CG	CA	GT	GC	AG	T	GC	AG	CG	T	CT	CG	T	CT	CG	T	1020	
Qy	1021	CC	GG	CC	CT	TA	TC	GG	AA	GT	CC	CC	CT	TC	CT	GG	CA	CG	TT	TC	GA	AG	CG	GT	TC	T	AG	CT	CG	CT	1080		
Db	1021	CC	GG	CC	CT	TA	TC	GG	AA	GT	CC	CC	CT	TC	CT	GG	CA	CG	TT	TC	GA	AG	CG	GT	TC	T	AG	CT	CG	CT	1080		
Qy	1081	GG	CA	T	GG	C	AG	GG	CT	TG	CC	CG	CA	CT	TG	CC	CG	CA	CT	TG	CC	CG	CA	CT	TG	CC	CG	CA	CT	TG	CC	1140	
Db	1081	GG	CA	T	GG	C	AG	GG	CT	TG	CC	CG	CA	CT	TG	CC	CG	CA	CT	TG	CC	CG	CA	CT	TG	CC	CG	CA	CT	TG	CC	1140	
Qy	1141	CC	CA	CT	CT	GC	CT	TG	CC	TAG	CG	GC	CA	CT	TAG	GG	GT	CAG	C	AG	C	CC	CT	TG	GG	CT	T	C	AC	CG	T	1200	
Db	1141	CC	CA	CT	CT	GC	CT	TG	CC	TAG	CG	GC	CA	CT	TAG	GG	GT	CAG	C	AG	C	CC	CT	TG	GG	CT	T	C	AC	CG	T	1200	
Qy	1201	GG	CC	CT	GG	T	GA	T	GA	T	GA	T	GA	T	GA	T	GA	T	GA	T	GA	T	GA	T	GA	T	GA	T	GA	T	GA	T	1260
Db	1201	GG	CC	CT	GG	T	GA	T	GA	T	GA	T	GA	T	GA	T	GA	T	GA	T	GA	T	GA	T	GA	T	GA	T	GA	T	GA	T	1260
Qy	1261	GT	AC	T	GT	GA	CT	TC	GG	CG	GC	GA	CT	TG	AG	CC	CG	TG	TG	GA	CT	TG	GC	CA	T	GT	GA	GG	CA	T	GT	1320	
Db	1261	GT	AC	T	GT	GA	CT	TC	GG	CG	GC	GA	CT	TG	AG	CC	CG	TG	TG	GA	CT	TG	GC	CA	T	GT	GA	GG	CA	T	GT	1320	
Qy	1321	CG	T	GG	CC	CT	CA	T	TC	CG	CAG	AG	GG	CT	CT	CT	AC	T	GT	CC	CG	T	GG	CT	T	CC	T	C	AG	CT	T	1380	
Db	1321	CG	T	GG	CC	CT	CA	T	TC	CG	CAG	AG	GG	CT	CT	CT	AC	T	GT	CC	CG	T	GG	CT	T	CC	T	C	AG	CT	T	1380	
Qy	1381	CG	CC	T	CCA	T	GT	CG	GG	CC	CT	T	CC	CT	GT	C	AG	CC	CG	AG	CG	CG	T	CA	AG	T	CT	GT	CG	T	CG	1440	
Db	1381	CG	CC	T	CCA	T	GT	CG	GG	CC	CT	T	CC	CT	GT	C	AG	CC	CG	AG	CG	CG	T	CA	AG	T	CT	GT	CG	T	CG	1440	
Qy	1441	GG	T	CG	TC	CG	CC	CT	GC	CT	CA	CC	CA	CT	GT	CT	GA	CT	GT	CT	GA	CT	GT	CT	GA	CT	GT	CT	GA	CT	GT	1500	
Db	1441	GG	T	CG	TC	CG	CC	CT	GC	CT	CA	CC	CA	CT	GT	CT	GA	CT	GT	CT	GA	CT	GT	CT	GA	CT	GT	CT	GA	CT	GT	1500	
Qy	1501	CC	GG	GA	T	GA	CT	TC	GG	CG	CT	T	CG	CC	CG	CG	AG	GG	GA	CT	CA	GG	CC	CT	T	AG	CT	AT	GC	T	1560		
Db	1501	CC	GG	GA	T	GA	CT	TC	GG	CG	CT	T	CG	CC	CG	CG	AG	GG	GA	CT	CA	GG	CC	CT	T	AG	CT	AT	GC	T	1560		
Qy	1561	TG	CG	CC	GG	G	AG	T	CG	GA	AG	AG	CT	CT	CT	GT	AT	TC	AC	CG	AG	CC	CT	T	AG	CT	T	CT	C	T	1620		
Db	1561	TG	CG	CC	GG	G	AG	T	CG	GA	AG	AG	CT	CT	CT	GT	AT	TC	AC	CG	AG	CC	CT	T	AG	CT	T	CT	C	T	1620		
Qy	1621	TG	AT	T	GG	A	T	CT	AT	T	CG	AA	G	TT	CT	GA	AG	CT	T	G	GA	A	CC	CT	T	AG	CT	AT	GC	T	1680		
Db	1621	TG	AT	T	GG	A	T	CT	AT	T	CG	AA	G	TT	CT	GA	AG	CT	T	G	GA	A	CC	CT	T	AG	CT	AT	GC	T	1680		
Qy	1681	TGG	CT	T	CC	CT	C	AG	T	GA	CC	CT	CA	T	CT	GT	C	AG	C	AG	CG	AG	GG	CC	CC	C	AG	GG	CT	GA	GG	1740	
Db	1681	TGG	CT	T	CC	CT	C	AG	T	GA	CC	CT	CA	T	CT	GT	C	AG	C	AG	CG	AG	GG	CC	CC	C	AG	GG	CT	GA	GG	1740	
Qy	1741	CAG	CA	T	T	G	T	AG	AG	C	C	AG	GG	AA	CC	AA	CT	T	T	G	GA	A	CC	CC	CA	CC	CT	T	CA	T	GA	1800	
Db	1741	CAG	CA	T	T	G	T	AG	AG	C	C	AG	GG	AA	CC	AA	CT	T	T	G	GA	A	CC	CC	CA	CC	CT	T	CA	T	GA	1800	
Qy	1801	AGA	A	CT	GT	CT	GA	GG	C	AG	GA	GT	CT	TA	CG	CC	AG	CA	AG	T	GA	GG	CT	T	GT	CA	GG	GG	GT	GA	1860		
Db	1801	AGA	A	CT	GT	CT	GA	GG	C	AG	GA	GT	CT	TA	CG	CC	AG	CA	AG	T	GA	GG	CT	T	GT	CA	GG	GG	GT	GA	1860		

QY	1861	CGGCTTT	CAGCCCTCTGGCTTTGGCCCTTTGCTTCA	CAGTGTAAATATCCCTCCCA	TCT	1920
DB	1861	CGGCTTT	CAGCCCTCTGGCTTTGGCCCTTTGCTTCA	CAGTGTAAATATCCCTCCCA	TCT	1920
QY	1921	TCCTTTCC	CCCTCTCTTCCCTTTCCCTCTCTCC	CCCTCGGTGAATGATGGCTTCTAAAA	1980	
DB	1921	TCTCTTCC	CCCTCTCTTCCCTTTCCCTCTCTCC	CCCTCGGTGAATGATGGCTTCTAAAA	1980	
QY	1981	CAAAATCA	ACCAAAACCTCAGCAGTGTGCATCTAT	AGCAGTGGCCCACTACCTGGCTCCA	2040	
DB	1981	CAAAATCA	ACCAAAACCTCAGCAGTGTGCATCTAT	AGCAGTGGCCCACTACCTGGCTCCA	2040	
QY	2041	CTGATCAC	CTCTCTGTGACCATCACCAACGGGTGC	CTTTGGCTGGCTTTTCCCTTG	2100	
DB	2041	CTGATCAC	CTCTCTGTGACCATCACCAACGGGTGC	CTTTGGCTGGCTTTTCCCTTG	2100	
QY	2101	GCCTTCCT	CAGCTTACCTTTGATCTAGCTGGGCTCTT	CTCTTGTCTGAAGCTGTGGAC	2160	
DB	2101	GCCTTCCT	CAGCTTACCTTTGATCTAGCTGGGCTCTT	CTCTTGTCTGAAGCTGTGGAC	2160	
QY	2161	CAGAGACCT	CGGACTTTTGTCTGCTTAAAGGAAATAG	GGGAAGTAAAGACAGTGAAGGGGT	2220	
DB	2161	CAGAGACCT	CGGACTTTTGTCTGCTTAAAGGAAATAG	GGGAAGTAAAGACAGTGAAGGGGT	2220	
QY	2221	GGAGGGTT	GTATCAGGGCACAGTGGACAGGAGACCT	CACAGAGAAAGCCCTGGAAGGTGA	2280	
DB	2221	GGAGGGTT	GTATCAGGGCACAGTGGACAGGAGACCT	CACAGAGAAAGCCCTGGAAGGTGA	2280	
QY	2281	TTTCCCGT	GTGACTCATGGATAGAGATACAAAATGT	GTGTCCATGTACATTAATCTTGACA	2340	
DB	2281	TTTCCCGT	GTGACTCATGGATAGAGATACAAAATGT	GTGTCCATGTACATTAATCTTGACA	2340	
QY	2341	TATGCCAT	GATAAAGACTTCCCTATTAAATAGCTTT	TGGAAGAGATTTAAAAA	2400	
DB	2341	TATGCCAT	GATAAAGACTTCCCTATTAAATAGCTTT	TGGAAGAGATTTAAAAA	2400	
QY	2401	AAAGGGCG	CGCGCTCTAGAGGATCCAAAGCTTACGT	ACGCTGCATGCGAGCTCATAGCTC	2460	
DB	2401	AAAGGGCG	CGCGCTCTAGAGGATCCAAAGCTTACGT	ACGCTGCATGCGAGCTCATAGCTC	2460	
QY	2461	TTCTATAG	TGTACCTAAATTCAAATT	2486		
DB	2461	TTCTATAG	TGTACCTAAATTCAAATT	2486		
RESULT 4						
ADQ76775						
ID	ADQ76775	standard; DNA; 3451 BP.				
XX	ADQ76775;					
AC						
XX						
DT	07-OCT-2004	(first entry)				
XX						
DE		Human LGR6 splice variant LGR6.2 nucleotide sequence.				
XX						
KW		Human; leucine-rich repeat-containing G-protein coupled receptor;				
KW		receptor; LGR6.2; cytosolic; immunomodulator; muscular-gen.;				
KW		gynaecological; antifertility; endocrine-gen.; anabolic; hypertensive;				
KW		vulnary; dermatological; gene therapy; gene; ds.				
XX						
OS		Homo sapiens.				
XX						
FH	Key	Location/Qualifiers				
FT	CDS	47..1093				
FT		/*tag= b				
FT		/product= "Human LGR6.2"				
FT	sig_peptide	47..103				
FT		/*tag= a				
FT	mat_peptide	104..1090				
FT		/*tag= c				
XX						
PN	WO2004058818-A2.					
XX						

PD	15-JUL-2004.
XX	
PF	23-DEC-2003; 2003WO-EF051091.
XX	
PR	26-DEC-2002; 2002US-0436001P.
XX	
PA	(ISTP) ARS APPLIED RES SYSTEMS HOLDING NV.
XX	
PI	Muda M;
XX	
XX	WPI; 2004-534126/51.
DR	
XX	
PT	Novel isolated leucine-rich repeat-containing G-protein coupled receptor-
PT	6 splicing variant 1 and 2 polypeptides, useful for treating or
PT	preventing endometriosis, female or male infertility, Addison's disease,
PT	Cushing's disease, cancers.
XX	
PS	Claim 1; SEQ ID NO 4; 85pp; English.
XX	
CC	The present nucleotide sequence encodes human LGR6.2, a novel splice
CC	variant of leucine-rich repeat-containing G-protein coupled receptor 6
CC	(LGR6-SV). It was initially identified as an expressed sequence tag
CC	clone. The LGR6.2 coding sequence consists of the same exons present in
CC	another splice variant, LGR6.1 ADO76773, plus an extended exon that
CC	introduces a stop codon resulting in truncation of the coding sequence,
CC	removing the transmembrane domains and the cytoplasmic C-terminal domain
CC	of LGR6.1. Thus, LGR6.2 is a secreted, probably soluble, N-terminal
CC	extracellular domain version of LGR6.1 and probably functions as an
CC	antagonist of the LGR6.1 signalling pathway. Tissue distribution studies
CC	indicate a high level of expression in ovary and pituitary tissues. The N
CC	-terminal extracellular domain of LGR6.2 can be used as an antagonist of
CC	the LGR6 signalling pathway. The invention provides LGR6.1 and LGR6.2
CC	polypeptides and nucleic acids, as well as selective binding agents,
CC	vectors, host cells and methods for producing LGR6-SV polypeptides. Also
CC	provided are pharmaceutical compositions and methods for the diagnosis,
CC	treatment, amelioration and/or prevention of diseases, disorders and
CC	conditions associated with LGR6-SV polypeptides. LGR6-SV polypeptides and
CC	nucleic acids can be used for diagnosing or treating: cancers such as
CC	hepatocellular carcinoma, adenomatous polyposis, malignant transformation
CC	of epithelial cells, colorectal cancer, breast cancer, ovarian cancer,
CC	etc.; cachexia and muscular dystrophy; miscarriage, endometriosis,
CC	uterine cancer, and female infertility; Cushing's disease and Addison's
CC	disease; male infertility and testicular cancer; leukaemia; and diseases
CC	and conditions affecting the kidney. They may also be useful for treating
CC	diseases and conditions involving cell proliferation and differentiation,
CC	such as tissue damage and degeneration, ageing, wound healing, cancer,
CC	hyperplasia and hypertrophy.
XX	
SQ	Sequence 3451 BP; 728 A; 1063 C; 920 G; 740 T; 0 U; 0 Other;
	Query Match 92.9%; Score 2309.8; DB 12; Length 3451;
	Best Local Similarity 99.7%; Pred. No. 0;
	Matches 2325; Conservative 0; Mismatches 7; Indels 1; Gaps 1;
Qy	71 CGGTCCGTGGAGCGGACCGAGGCTTCAGCCTGCCGGCTCATCCAGCCTCTTTGCTGCC 130
Db	1066 GGGGGCATGGAGCGGACCAGGCTTCAGCCTGCCGGCTCATCCAGCCTCTTTGCTGCC 1125
Qy	131 CTAGCGGCCTCCAACAACAACCGCATCTGGGAATTTGAGCTNGACACTTTCAGCACACTG 190
Db	1126 CTAGCGGCCTCCAACAACAACCGCATCTGGGAATTTGAGCT-GACACCTTTTCAGCACACTG 1184
Qy	191 AGCTCCCTCGAAGCCCTGGATCTTAGCTGGAAAGCCATCCGGTCCATCCACCTTGAGGCC 250
Db	1185 AGCTCCCTCGAAGCCCTGGATCTTAGCTGGAAAGCCATCCGGTCCATCCACCTTGAGGCC 1244
Qy	251 TTCTCCACCTTGCACTCCCTGGTCAAGCTGGACCTGACAGAACACACCACTGACCACTG 310
Db	1245 TTCTCCACCTTGCACTCCCTGGTCAAGCTGGACCTGACAGAACACCACTGACCACTG 1304
Qy	311 CCCTGCTGGACTTTGGGGCTTGATGCATCTGAAGCTCAAAGGAACCTTGCTCTCTCC 370
Db	1305 CCCTGCTGGACTTTGGGGCTTGATGCATCTGAAGCTCAAAGGAACCTTGCTCTCTCC 1364

QY 1451 CTGCTGCTGCTCAACCCACTGCTGCTGCTTCAACCCCACTTCCGGATGAC 1510
 Db 2445 CTGCTGCTGCTCAACCCACTGCTGCTGCTTCAACCCCACTTCCGGATGAC 2504
 QY 1511 CTTGGCGGCTTGGCGGCGGCGGAGCTCAGGGCCCTAGCCTATGCTGCGGCGGG 1570
 Db 2505 CTTGGCGGCTTGGCGGCGGCGGAGCTCAGGGCCCTAGCCTATGCTGCGGCGGG 2564
 QY 1571 GAGCTGGAGAGAGCTCTGTGATTTCTACCCAGGCGCTGGTATGCTTCTGATGTGAT 1630
 Db 2565 GAGCTGGAGAGAGCTCTGTGATTTCTACCCAGGCGCTGGTATGCTTCTGATGTGAT 2624
 QY 1631 CTATTCTGGAAGCTTCTGAAGCTGGGCGGCGGCTGGGCTGGAGACCTATGCTTCCCC 1690
 Db 2625 CTATTCTGGAAGCTTCTGAAGCTGGGCGGCGGCTGGGCTGGAGACCTATGCTTCCCC 2684
 QY 1691 TCAGTGACCTCATCTCTGTGAGCAGCAGGGGCGGCGGCTGGGCTGGAGGAGCATTGT 1750
 Db 2685 TCAGTGACCTCATCTCTGTGAGCAGCAGGGGCGGCGGCTGGGCTGGAGGAGCATTGT 2744
 QY 1751 GTAGAGCAGAGGGAGCACTTTGGGAACCCCAACCTCCATGGATGGAGACTGCTG 1810
 Db 2745 GTAGAGCAGAGGGAGCACTTTGGGAACCCCAACCTCCATGGATGGAGACTGCTG 2804
 QY 1811 CTAGGGCAGAGGGATCTACGCCAGGCTGGAGGCTTGTCAAGGGGTGGCGGCTTTCAG 1870
 Db 2805 CTAGGGCAGAGGGATCTACGCCAGGCTGGAGGCTTGTCAAGGGGTGGCGGCTTTCAG 2864
 QY 1871 CCCTCTGCTGGCTTGGCTTTCACACCTGTAAATATCCCTCCCATTTCTTCTTCCCC 1930
 Db 2865 CCCTCTGCTGGCTTGGCTTTCACACCTGTAAATATCCCTCCCATTTCTTCTTCCCC 2924
 QY 1931 TCTCTTCCCTTCTCTCCCTCGTGAATGCTGCTCTTCAAAACAAATACAAAC 1990
 Db 2925 TCTCTTCCCTTCTCTCCCTCGTGAATGCTGCTCTTCAAAACAAATACAAAC 2984
 QY 1991 CAAAACTCAGCAGTGTGATCTATAGCAGATGGCCAGTACTGCTGCTCACTGATCACT 2050
 Db 2985 CAAAACTCAGCAGTGTGATCTATAGCAGATGGCCAGTACTGCTGCTCACTGATCACT 3044
 QY 2051 CTCTCTGTGACCATACCAACAGGCTGCTTGGCTGCTTCCCTTGGCTTCCCTCA 2110
 Db 3045 CTCTCTGTGACCATACCAACAGGCTGCTTGGCTGCTTCCCTTGGCTTCCCTCA 3104
 QY 2111 GCTTCACCTTCACTAGGCTTCTTCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2170
 Db 3105 GCTTCACCTTCACTAGGCTTCTTCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3164
 QY 2171 GACTTTTGTCTGCTTAAAGGAAATGAGGAAATGAGGAAATGAGGAAATGAGGAAATG 2230
 Db 3165 GACTTTTGTCTGCTTAAAGGAAATGAGGAAATGAGGAAATGAGGAAATGAGGAAATG 3224
 QY 2231 TCAGGGCAGCTGAGCAGGGAGACCTCAACAGAAAGGCTGGAAGGTTGATTTCCCGTGT 2290
 Db 3225 TCAGGGCAGCTGAGCAGGGAGACCTCAACAGAAAGGCTGGAAGGTTGATTTCCCGTGT 3284
 QY 2291 GACTCATGGATAGGATACAAATGTTCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2350
 Db 3285 GACTCATGGATAGGATACAAATGTTCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3344
 QY 2351 ATAAAGACTTCTCTATTAATAAAGCTTTGGAGAGATTTAAAAAATAAATAAATAAATA 2403
 Db 3345 ATAAAGACTTCTCTATTAATAAAGCTTTGGAGAGATTTAAAAAATAAATAAATAAATA 3397

RESULT 5
 AAF98722 standard; DNA; 2322 BP.
 ID AAF98722
 XX
 AC AAF98722;
 XX
 DT 02-JUL-2001 (first entry)
 XX

DE Human late stage ovarian tumour polynucleotide marker 37.
 XX Human; ovarian cancer; identification; detection; characterisation;
 KW tumour; kinase; marker; cytostatic; antisense gene therapy; ds.
 OS Homo sapiens.
 PN WO200118542-A2.
 PD 15-MAR-2001.
 XX 01-SEP-2000; 2000WO-US024199.
 XX 03-SEP-1999; 99US-0152547P.
 PR 16-MAR-2000; 2000US-0190347P.
 PR 21-MAR-2000; 2000US-0191321P.
 PR 31-MAY-2000; 2000US-0208382P.
 PR 20-JUL-2000; 2000US-00220467.
 XX (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.
 PA Lee J, Thompson P, Lillie J;
 FI WPI; 2001-211428/21.
 DR Detection, assessment, prevention and therapy of ovarian cancer.
 PT comprises detecting changes in the expression of a variety of markers.
 XX Claim 1; Page 1188-1189; 1198pp; English.
 PS The present invention describes a method for assessing whether a patient
 XX is afflicted with ovarian cancer by comparing: (1) the expression of a
 CC marker (1) (see AAF98594 to AAF98730), in a patient sample; and (2) the
 CC normal level of expression of (1) in a control non-ovarian cancer sample,
 CC where a significant difference between the level of expression in (a) and
 CC (b) is an indication that the patient is afflicted with ovarian cancer.
 CC (b) have cytostatic activities and can be used in antisense gene therapy.
 CC The method, compositions and kits from the present invention can be used
 CC for: (1) assessing and treating ovarian cancer; (2) making isolated
 CC hybridoma, which produces an antibody useful for ovarian cancer
 CC assessment; and (3) inhibiting ovarian cancer in a patient. AAF98573 to
 CC AAF98593 represent human kinase marker primers and probes which are used
 CC in the exemplification of the present invention
 XX Sequence 2322 BP; 407 A; 733 C; 652 G; 530 T; 0 U; 0 Other;
 SQ Query Match 92.9%; Score 2309.4; DB 5; Length 2322;
 Best Local Similarity 99.9%; Pred. No. 0;
 Matches 2321; Conservative 0; Mismatches 1; Indels 1; Gaps 1;
 QY 67 CCAGCGTCCGTGGAGCGGAGCCAGGGTCTGAGCCTGCGGCTCATCCAGCCTCTTGC 126
 Db 1 CCAGCGTCCGTGGAGCGGAGCCAGGGTCTGAGCCTGCGGCTCATCCAGCCTCTTGC 60
 QY 127 TGCCCTAGCGGCTCCCAACCAACCGCATCTGGGAAATTTGGAGCTGACACCTTCAGCCA 186
 Db 61 TGCCCTAGCGGCTCCCAACCAACCGCATCTGGGAAATTTGGAGCT-GACACCTTCAGCCA 119
 QY 187 GCTGAGCTCCCTCAAGCCCTGGATCTTAGCTGGAACGCCATCGGTCCATFCAACCTGA 246
 Db 120 GCTGAGCTCCCTCAAGCCCTGGATCTTAGCTGGAACGCCATCGGTCCATFCAACCTGA 179
 QY 247 GGCTTTCTCCACCTGACCTCCCTGCTCAAGCTGGACCTGACAGACACAGCTGACAC 306
 Db 180 GGCTTTCTCCACCTGACCTCCCTGCTCAAGCTGGACCTGACAGACACAGCTGACAC 239
 QY 307 ACTGCCCTGGCTGACTTTGGGGCTTGTAGTCTGAAGCTCAAGGAAACCTTGTCT 366
 Db 240 ACTGCCCTGGCTGACTTTGGGGCTTGTAGTCTGAAGCTCAAGGAAACCTTGTCT 299
 QY 367 CTCCCAGGCTTCTCCAAAGGACAGTTTCCAAATCTGAGGATCTTGAGGCTTATGC 426
 Db 300 CTCCCAGGCTTCTCCAAAGGACAGTTTCCAAATCTGAGGATCTTGAGGCTTATGC 359

```
QY 427 CTACAGTGTCTCCCTATGGGATGTGTCCAGCTTCTTCAAGGCTCTGGGCTCTCTTGC 486
Db |
QY 360 CTACAGTGTCTCCCTATGGGATGTGTCCAGCTTCTTCAAGGCTCTGGGCTCTTGC 419
Db |
QY 487 GGCTGAAGACCTTCACTTGTATGATGAGGAGTCTTCAAAAGGCGCCCTGGGCTCTCTTGC 546
Db |
QY 420 GGCTGAAGACCTTCACTTGTATGATGAGGAGTCTTCAAAAGGCGCCCTGGGCTCTCTTGC 479
QY 547 CAGACAAGCAGAAACCACTATGACACGAGCCTGGATGAGCTTCCAGCTGGAGATGAGGA 606
Db |
QY 480 CAGACAAGCAGAAACCACTATGACACGAGCCTGGATGAGCTTCCAGCTGGAGATGAGGA 539
QY 607 CTAAAGCCACACCCCAAGTGTCCAGTGTAGCCCTACTCCAGGCCCTTCAAGCCCTGTGA 666
Db |
QY 540 CTAAAGCCACACCCCAAGTGTCCAGTGTAGCCCTACTCCAGGCCCTTCAAGCCCTGTGA 599
QY 667 GTACCTCTTTGAAAGCTGGGCGATCCGCTTGGCCGTGTGGCCATCGTGTCTCCGT 726
Db |
QY 600 GTACCTCTTTGAAAGCTGGGCGATCCGCTTGGCCGTGTGGCCATCGTGTCTCCGT 659
QY 727 GCTCTGCAATGAGCTGTGTCTGACACGCTGTTCGCTGGCGGCTGCCCCCTGCCGCC 786
Db |
QY 660 GCTCTGCAATGAGCTGTGTCTGACACGCTGTTCGCTGGCGGCTGCCCCCTGCCGCC 719
QY 787 GGTCAAGTTTGTGTAGTGTGGATTGSCAGGGCGCCAAACCTTTGACTGGCATTTCTGTGG 846
Db |
QY 720 GGTCAAGTTTGTGTAGTGTGGATTGSCAGGGCGCCAAACCTTTGACTGGCATTTCTGTGG 779
QY 847 CTTCTAGCTCAGTGTGATGCCCTGACCTTTGGTGTGATGTCTGTAGTACGAGGCCGCTG 906
Db |
QY 780 CTTCTAGCTCAGTGTGATGCCCTGACCTTTGGTGTGATGTCTGTAGTACGAGGCCGCTG 839
QY 907 GGAGACGGGCTTAGCTGCGGGCCACTGGCTTCTGCGAGTACTTGGGTGGAGGCATC 966
Db |
QY 840 GGAGACGGGCTTAGCTGCGGGCCACTGGCTTCTGCGAGTACTTGGGTGGAGGCATC 899
QY 967 GGTGTGTGTCTCACTCTGGCGCAGTGTGACAGCGTCTCCGTCTCTGTGTCCGGGC 1026
Db |
QY 900 GGTGTGTGTCTCACTCTGGCGCAGTGTGACAGCGTCTCCGTCTCTGTGTCCGGGC 959
QY 1027 CTATGGGAAGTCCCTCTCTGGGAGAGGTTGCGAGCAGGGTCTTAGGCTGTGCGACT 1086
Db |
QY 960 CTATGGGAAGTCCCTCTCTGGGAGAGGTTGCGAGCAGGGTCTTAGGCTGTGCGACT 1019
QY 1087 GGCAGGGCTGCGCGCCGACTGCCCCCTGCGCTCTAGTGGGAGAAATACGGGGCTCCCACT 1146
Db |
QY 1020 GGCAGGGCTGCGCGCCGACTGCCCCCTGCGCTCTAGTGGGAGAAATACGGGGCTCCCACT 1079
QY 1147 CTGCTGCTCCCTACGCGCCACCTGAGGGTTCAGCCAGCAGCCCTTGGGCTTTCACCGTGGCCCT 1206
Db |
QY 1080 CTGCTGCTCCCTACGCGCCACCTGAGGGTTCAGCCAGCAGCCCTTGGGCTTTCACCGTGGCCCT 1139
QY 1207 GGTGATGATGAATCTCTTCTGTTTCTGGTGTGCGTGGCGGCTACATCAAACTGTACTG 1266
Db |
QY 1140 GGTGATGATGAATCTCTTCTGTTTCTGGTGTGCGTGGCGGCTACATCAAACTGTACTG 1199
QY 1267 TGACTGCGCGGGGCGACTTTGAGGCGGTGTGGACTGCGCCATGGTGTAGGACAGTGGC 1326
Db |
QY 1200 TGACTGCGCGGGGCGACTTTGAGGCGGTGTGGACTGCGCCATGGTGTAGGACAGTGGC 1259
QY 1327 CTGGCTCATCTTTCGAGACGGGCTCTCTACTGTCTCCGTGGCTTTCCTCAGCTTCGCGCTC 1386
Db |
QY 1260 CTGGCTCATCTTTCGAGACGGGCTCTCTACTGTCTCCGTGGCTTTCCTCAGCTTCGCGCTC 1319
QY 1387 CATGTGGGCTCTTTCCTGTTCAGCCCGAGGCGGTCAAGTGTGTCTGTGTGGTGTCT 1446
Db |
QY 1320 CATGTGGGCTCTTTCCTGTTCAGCCCGAGGCGGTCAAGTGTGTCTGTGTGGTGTCT 1379
QY 1447 GCGGCTGCTGCTGCTCAACCCACTGTGTACCTGTCTTCAACCCCACTTCCGGGA 1506
Db |
QY 1380 GCGGCTGCTGCTGCTCAACCCCACTGTGTACCTGTCTTCAACCCCACTTCCGGGA 1439
```

```
QY 1507 TGACCTTCGGCGGCTTCGGCCCCGCGCAGGGAGCTCAGGGCCCCCTAGCTATGTCTGCGGC 1566
Db |
QY 1440 TGACCTTCGGCGGCTTCGGCCCCGCGCAGGGAGCTCAGGGCCCCCTAGCTATGTCTGCGGC 1499
QY 1567 CGGGAGCTTGAGAAAGAGCTCTGTGATTTACCCAGGCCCCCTGTGTAGCTTCTCTGTATGT 1626
Db |
QY 1500 CGGGAGCTTGAGAAAGAGCTCTGTGATTTACCCAGGCCCCCTGTGTAGCTTCTCTGTATGT 1559
QY 1627 GGATCTCATTTCTGGAAGCTTCTGAAAGCTTGGGCGGCCCCCTGGGCTGGAGACCTATGGCTT 1686
Db |
QY 1560 GGATCTCATTTCTGGAAGCTTCTGAAAGCTTGGGCGGCCCCCTGGGCTGGAGACCTATGGCTT 1619
QY 1687 CCCCTCAGTGACCTCATCTCTGTACGACGAGGGGCCCCCAGGCTGGAGGGCAGGCA 1746
Db |
QY 1620 CCCCTCAGTGACCTCATCTCTGTACGACGAGGGGCCCCCAGGCTGGAGGGCAGGCA 1679
QY 1747 TTGTGTAGAGCCAGAGGGGAACCACTTTGGGAAACCCCAACCCCTCCATGGATGGAGACT 1806
Db |
QY 1680 TTGTGTAGAGCCAGAGGGGAACCACTTTGGGAAACCCCAACCCCTCCATGGATGGAGACT 1739
QY 1807 GCTGTGAGGCGCAGAGGGATCTAGCCAGCAGGTGGAGGCTTGTTCAGGGGCTGGGGCTT 1866
Db |
QY 1740 GCTGTGAGGCGCAGAGGGATCTAGCCAGCAGGTGGAGGCTTGTTCAGGGGCTGGGGCTT 1799
QY 1867 TCAGCCCTCTGGCTTGGCTTTGCTTTCACAGTGTAAATATCCCTCCCACTTCTCTCTT 1926
Db |
QY 1800 TCAGCCCTCTGGCTTGGCTTTTGTCTTACACGTTGTAATATCCCTCCCACTTCTCTCTT 1859
QY 1927 CCCCTCTCTCCCTTCTCTCTCCCTCGGTGAATGATGGCTGTCTTAAACAAATA 1986
Db |
QY 1860 CCCCTCTCTCCCTTCTCTCTCCCTCGGTGAATGATGGCTGTCTTAAACAAATA 1919
QY 1987 CAACCAAACTCAGCAGTGTGATCTATAGCAGGATGGCCAGTACCTGGCTCCACTGATC 2046
Db |
QY 1920 CAACCAAACTCAGCAGTGTGATCTATAGCAGGATGGCCAGTACCTGGCTCCACTGATC 1979
QY 2047 ACCTCTCTCTGTGACCATCACCAACGGGTGCTTCTTGGCTGGCTTTCCTTGGCCCTTC 2106
Db |
QY 1980 ACCTCTCTCTGTGACCATCACCAACGGGTGCTTCTTGGCTGGCTTTCCTTGGCCCTTC 2039
QY 2107 CTGAGCTTCACCTTGATCTAGGGCTCTTCTTGTCTCATGTCTGAGCTGTGGACCAAGA 2166
Db |
QY 2040 CTGAGCTTCACCTTGATCTAGGGCTCTTCTTGTCTCATGTCTGAGCTGTGGACCAAGA 2099
QY 2167 CCTGACCTTTTGTCTGCTTAAAGGAAATGAGGGAAGTAAAGACAGTGAAGGGTGGAGGG 2226
Db |
QY 2100 CCTGACCTTTTGTCTGCTTAAAGGAAATGAGGGAAGTAAAGACAGTGAAGGGTGGAGGG 2159
QY 2227 TTGATCAGGGCAGTGTGACAGAGGAGACCTCACAGAGAAAGGCTTGGAGGTGATTTCCC 2286
Db |
QY 2160 TTGATCAGGGCAGTGTGACAGAGGAGACCTCACAGAGAAAGGCTTGGAGGTGATTTCCC 2219
QY 2287 GTGTGACTCATGTGATAGGATACAAATGTGTTCATGTACATTAATTTGACATATGCC 2346
Db |
QY 2220 GTGTGACTCATGTGATAGGATACAAATGTGTTCATGTACATTAATTTGACATATGCC 2279
QY 2347 ATGCATAAAGACTTCTTAAATAAAGCTTTGGAAGAGATTA 2389
Db |
QY 2280 ATGCATAAAGACTTCTTAAATAAAGCTTTGGAAGAGATTA 2322
```

RESULT 6

AAI67924
ID AAI67924 standard; cDNA; 2711 BP.

XX AAI67924;

XX AC

DT 13-MAR-2002 (first entry)

XX

DE Partial cDNA of human LGR6 polypeptide (clone fahr).

XX

KW Large G-protein coupled receptor 6; LGR6; GPCR; immunomodulator;

KW anti-HIV; cytostatic; nootropic; neuroprotective; antiparkinsonian;

Db 1701 GGGCCCTTCCCTGTGACGCCGAGCGGTCAAGTCTGTCTGCTGGTGGTGCCTGCCCT 1760
Qy GCCTGCTGCTCAACCCACCTGCTGTACTGTCTTCAACCCCACTTCCGGGATGACT 1512
Db GCCTGCTGCTCAACCCACCTGCTGTACTGTCTTCAACCCCACTTCCGGGATGACT 1820
Qy TCGGGGCTTGGCCCGCGCAGGAGGACTCAGGGCCCTAGCCTATGCTGCGGCGGGGA 1572
Db TCGGGGCTTGGCCCGCGCAGGAGGACTCAGGGCCCTAGCCTATGCTGCGGCGGGGA 1880
Qy GCTGAGAGAGCTCTCTGTGATTTACCCAGGCGCTGTAGCCTTCTGTGATGGATCT 1632
Db GCTGAGAGAGCTCTCTGTGATTTACCCAGGCGCTGTAGCCTTCTGTGATGGATCT 1940
Qy CATCTGGAAGCTTCTGAAGCTTGGCGGCCCTTGGCTGGAGACCTATGCTTCCCTCTC 1692
Db CATCTGGAAGCTTCTGAAGCTTGGCGGCCCTTGGCTGGAGACCTATGCTTCCCTCTC 2000
Qy AGTGACCTCATCTCTGTGACGACCCAGGGGCCCTAGGCTGGAGGCGCCTATGCTGT 1752
Db AGTGACCTCATCTCTGTGACGACCCAGGGGCCCTAGGCTGGAGGCGCCTATGCTGT 2060
Qy AGAGCCAGAGGGGACCACTTTGGGAACCCCAACCTCCATGGATGGAGAACTGCTGCT 1812
Db AGAGCCAGAGGGGACCACTTTGGGAACCCCAACCTCCATGGATGGAGAACTGCTGCT 2120
Qy GAGGCGAGAGGGATACGCCAGCAGGTGGAGGCTTGTGAGGGGTGGCGCTTTCAGCC 1872
Db GAGGCGAGAGGGATACGCCAGCAGGTGGAGGCTTGTGAGGGGTGGCGCTTTCAGCC 2180
Qy CTCTGGCTTGGCTTGTCTTCAACCTGTAATATCCCTCCCACTTCTCTTCCCTCTC 1932
Db CTCTGGCTTGGCTTGTCTTCAACCTGTAATATCCCTCCCACTTCTCTTCCCTCTC 2240
Qy TCTTCCCTTCTCTCTCCCTCGGTGAATGATGGCTGCTTCTTAAACAATAACAACA 1992
Db TCTTCCCTTCTCTCTCCCTCGGTGAATGATGGCTGCTTCTTAAACAATAACAACA 2300
Qy AAATCAGCAGTGTGATCTATAGCAGGATGCCCAGTACCTGGCTCCACTGATCACCCT 2052
Db AAATCAGCAGTGTGATCTATAGCAGGATGCCCAGTACCTGGCTCCACTGATCACCCT 2360
Qy CTCCTGTGACCATCACCAACGGGTGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCT 2112
Db CTCCTGTGACCATCACCAACGGGTGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCT 2420
Qy TTCACCTTGATCTGGGCTTCTCTTGTCTATGCTGAGCTGTGGACGAGACCTGGA 2172
Db TTCACCTTGATCTGGGCTTCTCTTGTCTATGCTGAGCTGTGGACGAGACCTGGA 2480
Qy CTTTGTCTGCTTAAGGGAATGAGGGAAGTAAAGACAGTGAAGGGGTGGAGGTTGATC 2232
Db CTTTGTCTGCTTAAGGGAATGAGGGAAGTAAAGACAGTGAAGGGGTGGAGGTTGATC 2540
Qy AGGCGACAGTGGACGAGGAGACCTCACAGAGAAAGGCTGGAAGTGATTTCCCGTGA 2292
Db AGGCGACAGTGGACGAGGAGACCTCACAGAGAAAGGCTGGAAGTGATTTCCCGTGA 2600
Qy CTCATGATAGATACAAAATGTGTTCCATGATACCAATTAATCTTGACATATGCCATGAT 2352
Db CTCATGATAGATACAAAATGTGTTCCATGATACCAATTAATCTTGACATATGCCATGAT 2660
Qy AAAGACTTCTTAAATAAGCTTTGGAAGAGATTAATAAAAAA 2403
Db AAAGACTTCTTAAATAAGCTTTGGAAGAGATTAATAAAAAA 2711

RESULT 7
ID ADN02245 standard; cDNA; 2711 BP.
XX
AC ADN02245;

XX 17-JUN-2004 (first entry)
DT Human partial large G protein-coupled receptor, LGR6, cDNA #2.
XX
DE Human; ss; large G protein-coupled receptor 6; LGR6; heart failure;
KW ischaemic heart disease; myocardial infarction; hypertension;
KW pericarditis; atherosclerosis; hepatic failure; viral hepatitis; cancer;
KW intracerebral haemorrhage; brain abscess; Alzheimer's disease;
KW Pick disease; Parkinson's disease; Huntington's disease;
KW amyotrophic lateral sclerosis; inflammatory bowel disease;
KW Crohn's disease; benign prostatic hypertrophy;
KW systemic lupus erythematosus.
XX
OS Homo sapiens.
XX
XX Key Location/Qualifiers
CDS 1..2211 a
FT /*tag=
FT /partial
FT /note= "No start codon shown"
FT /product= "Partial LGR6"
XX
XX US2003166047-A1.
XX
XX 04-SEP-2003.
XX
XX 08-MAY-2001; 2001US-00851595.
XX
XX 06-MAY-1999; 39US-0132896P.
XX 08-MAY-2000; 2000US-00566588.
XX (MILL-) MILLENNIUM PHARM INC.
XX
XX Gu W;
XX
XX WPI; 2003-898067/82.
XX P-PSDB; ADN02246.
XX
XX New isolated LGR6 nucleic acids and polypeptides, useful for diagnosing
PT and treating LGR6-mediated disorders, such as myocardial infarction,
PT hyperextension, atherosclerosis, viral hepatitis, cancer and/or Alzheimer's
PT disease.
XX
XX Claim 1; SEQ ID NO 7; 145pp; English.
XX
XX The invention relates to an isolated LGR6 (large G protein-coupled
CC receptor) nucleic acid molecule comprising the mouse and human cDNA
CC sequences appearing as ADN02239, ADN02241, ADN02242, ADN02244, ADN02245,
CC ADN02247, ADN02248 and ADN02250. Also included are a host cell or a non-
CC human mammalian host cell containing an LGR6 nucleic acid (and producing
CC LGR6), an isolated LGR6 polypeptide appearing as ADN02240, ADN02243, a
CC ADN02246 and ADN02249, an antibody which selectively binds to LGR6, a
CC method for detecting the presence of LGR6 in a sample, a kit (comprising
CC a compound which selectively binds to LGR6, or hybridises to LGR6 nucleic
CC acid, and instructions for use), a method for detecting LGR6 nucleic acid
CC in a sample, a method for identifying a compound which binds to LGR6, a
CC method for modulating the activity of LGR6 and a method for identifying a
CC compound that modulates the activity of LGR6. The methods and
CC compositions of the present invention are useful for diagnosing and
CC treating large G-protein coupled receptor 6 (LGR6)-mediated or -related
CC disorders, such as heart failure, ischaemic heart disease, myocardial
CC infarction, hypertension, pericarditis, atherosclerosis, hepatic failure,
CC viral hepatitis, cancer, intracerebral haemorrhage, brain abscess,
CC Alzheimer's disease, Pick disease, Parkinson's disease, Huntington's
CC disease, amyotrophic lateral sclerosis, inflammatory bowel disease,
CC Crohn's disease, benign prostatic hypertrophy and systemic lupus
CC erythematosus. The present sequence is a partial human cDNA for LGR6.
XX
XX Sequence 2711 BP; 530 A; 837 C; 740 G; 604 T; 0 U; 0 Other;
XX

Query Match 90.8%; Score 2257.4; DB 11; Length 2711;
Best Local Similarity 99.9%; Pred. No. 0;

Matches 2269; Conservative 0; Mismatches 1; Indels 1; Gaps 1;			
Qy	133	AGCGCCTCCACACAAACCGCATCTGGGAAATTTGGAGCTNGACACACCTTCAGCCAGCTGAG	192
Db	442	ATCGGCCCTCCACACAAACCGCATCTGGGAAATTTGGAGCT-GACACCTTTCAGCCAGCTGAG	500
Qy	193	CTCCCTGCAAGCCCTTGGATCTTTAGCTGGAAAGCCATCCGGTCCATCCACCCCTGAGGCCCTT	252
Db	501	CTCCCTGCAAGCCCTTGGATCTTTAGCTGGAAAGCCATCCGGTCCATCCACCCCTGAGGCCCTT	560
Qy	253	CTCCACCTTGCACTCCCTGGTCAAGCTGGACCTTGACAGACAAACCAAGCTGACCACTGCC	312
Db	561	CTCCACCTTGCACTCCCTGGTCAAGCTGGACCTTGACAGACAAACCAAGCTGACCACTGCC	620
Qy	313	CCTGGCTGGACTTGGGGCTTGATGCACTGAAGCTCAAGGAAACCTTGCTCTCTCCCA	372
Db	621	CCTGGCTGGACTTGGGGCTTGATGCACTGAAGCTCAAGGAAACCTTGCTCTCTCCCA	680
Qy	373	GGCCTTCTCCAAAGACAGTTTCCCAAAACTGAGGATCTCGAGGTGCCCTTATGCTACCA	432
Db	681	GGCCTTCTCCAAAGACAGTTTCCCAAAACTGAGGATCTCGAGGTGCCCTTATGCTACCA	740
Qy	433	GTGCTGCTCCCTATGGGATGTGTCAGCTTCTTCAAGGCCCTCTGGGCAGTGGGAGGTGA	492
Db	741	GTGCTGCTCCCTATGGGATGTGTCAGCTTCTTCAAGGCCCTCTGGGCAGTGGGAGGTGA	800
Qy	493	AGACCTTCACCTTGATGATGAGGAGTCTTCAAAAAGGCCCTTGGCCCTCTTGGCAGACA	552
Db	801	AGACCTTCACCTTGATGATGAGGAGTCTTCAAAAAGGCCCTTGGCCCTCTTGGCAGACA	860
Qy	553	AGCAGAGAACCACTATGACACAGGACCTGGATGAGCTCCAGCTGGAGATGGAGGACTCAAA	612
Db	861	AGCAGAGAACCACTATGACACAGGACCTGGATGAGCTCCAGCTGGAGATGGAGGACTCAAA	920
Qy	613	GCCACACCCCAAGTGTCCAGTGTAGCCCTTACTCCAGGCCCTTCAAGCCCTGTGAGTACCT	672
Db	921	GCCACACCCCAAGTGTCCAGTGTAGCCCTTACTCCAGGCCCTTCAAGCCCTGTGAGTACCT	980
Qy	673	CTTTGAAAGCTGGGGCATCCGCTTGGCGTGTGGGCCATCGTGTGCTCTCGGTGCTGTG	732
Db	981	CTTTGAAAGCTGGGGCATCCGCTTGGCGTGTGGGCCATCGTGTGCTCTCGGTGCTGTG	1040
Qy	733	CAATGGACTGGTGTGCTGACCGTGTTCGCTGGGGGCTTCCGCCCTTGGCCCGGTCAA	792
Db	1041	CAATGGACTGGTGTGCTGACCGTGTTCGCTGGGGGCTTCCGCCCTTGGCCCGGTCAA	1100
Qy	793	GTTTGTGTAGTGTGCGATGTGAGCGCGCAACACCTTGACTGGCAATTCCTGTGGCCTTCT	852
Db	1101	GTTTGTGTAGTGTGCGATGTGAGCGCGCAACACCTTGACTGGCAATTCCTGTGGCCTTCT	1160
Qy	853	AGCCTCAGTCGATGCCCTGACCTTTGGTCACTTCTTGAGTACGGAGCCGCTGGGAGAC	912
Db	1161	AGCCTCAGTCGATGCCCTGACCTTTGGTCACTTCTTGAGTACGGAGCCGCTGGGAGAC	1220
Qy	913	GGGCTAGGCTGCGGGCCACTGGCTTCTCGCAGTACTTGGGTGGAGGCATCGGTGCT	972
Db	1221	GGGCTAGGCTGCGGGCCACTGGCTTCTCGCAGTACTTGGGTGGAGGCATCGGTGCT	1280
Qy	973	GCTGCTCACTGTCGCGCAGTGCAGTGCAGCGTCTCCGTCCTCTGTGTCCGGGCCCTATGG	1032
Db	1281	GCTGCTCACTGTCGCGCAGTGCAGTGCAGCGTCTCCGTCCTCTGTGTCCGGGCCCTATGG	1340
Qy	1033	GAAGTCCCTCTCCCTGGGACGCTTGGAGCAGGGGTCTTAGGCTGGCTGGACCTGGCAGG	1092
Db	1341	GAAGTCCCTCTCCCTGGGACGCTTGGAGCAGGGGTCTTAGGCTGGCTGGCAGCTGGCAGG	1400
Qy	1093	GCTGGCGCGCACTGGCCCTGGCTCACTGGGAGAAATACGGGGCCCTCCCACTCTGCCT	1152
Db	1401	GCTGGCGCGCACTGGCCCTGGCTCACTGGGAGAAATACGGGGCCCTCCCACTCTGCCT	1460
Qy	1153	GCCCTACGCGCACTGAGGGTACGCCAGCAGCCCTGGGCTTACCGGTGGCCCTGGTGTAT	1212
Db	1461	GCCCTACGCGCACTGAGGGTACGCCAGCAGCCCTGGGCTTACCGGTGGCCCTGGTGTAT	1520

Qy	1213	GATGAACCTCCTTCTGTTTTCTGTGTCGTGGCGCGGTCCCTACATCAAACTGTACTGTGACCT	1272
Db	1521	GATGAACCTCCTTCTGTTTTCTGTGTCGTGGCGCGGTCCCTACATCAAACTGTACTGTGACCT	1580
Qy	1273	GCOCGGGGGACATTTTGGAGCCGTGTGGGACTGCGCCATGGTGAAGGACGCTGGCCTGGCT	1332
Db	1581	GCOCGGGGGACATTTTGGAGCCGTGTGGGACTGCGCCATGGTGAAGGACGCTGGCCTGGCT	1640
Qy	1333	CATCTTTCGACAGACGGGCTCCTCTACTGTCCGTGGCCCTTCTCAGCTTCGGCTCAGTCT	1392
Db	1641	CATCTTTCGACAGACGGGCTCCTCTACTGTCCGTGGCCCTTCTCAGCTTCGGCTCAGTCT	1700
Qy	1393	GGGCTCCTTCCCTGCTACGCCCGGAGCCGTCAAGTCTGCTGCTGCTGCTGCTGCTGCTGCT	1452
Db	1701	GGGCTCCTTCCCTGCTACGCCCGGAGCCGTCAAGTCTGCTGCTGCTGCTGCTGCTGCTGCT	1760
Qy	1453	GCCTGCTGCTCTCAACCCACTGTCTGTACTGCTCTTCAACCCCCCACTTCGCGGATGACCT	1512
Db	1761	GCCTGCTGCTCTCAACCCACTGTCTGTACTGCTCTTCAACCCCCCACTTCGCGGATGACCT	1820
Qy	1513	TCGGCGGCTTCGGCCCGCGGAGGAGACTCAGGGCCCTAGCCTATGCTGCGGCGGGGA	1572
Db	1821	TCGGCGGCTTCGGCCCGCGGAGGAGACTCAGGGCCCTAGCCTATGCTGCGGCGGGGA	1880
Qy	1573	GCTGGAGAGAGCTCCTGTGATTTACCCAGGCCCTGCTAGCCTTCTCTGATGTGGATCT	1632
Db	1881	GCTGGAGAGAGCTCCTGTGATTTACCCAGGCCCTGCTAGCCTTCTCTGATGTGGATCT	1940
Qy	1633	CATTCTGGAAGCTTCTGAAGCTGGGGCGGCCCTTGGGCTGGAGACCTATGGCTTCCCTC	1692
Db	1941	CATTCTGGAAGCTTCTGAAGCTGGGGCGGCCCTTGGGCTGGAGACCTATGGCTTCCCTC	2000
Qy	1693	AGTGAACCTCATCTCTGTGACGAGCCAGGGGCCCCCAGGCTGGAGGCGACATTGTGT	1752
Db	2001	AGTGAACCTCATCTCTGTGACGAGCCAGGGGCCCCCAGGCTGGAGGCGACATTGTGT	2060
Qy	1753	AGAGCCAGAGGGGAAACCACTTTTGGGAACCCCAACCTCCATGATGATGAGAACTGCTGCT	1812
Db	2061	AGAGCCAGAGGGGAAACCACTTTTGGGAACCCCAACCTCCATGATGATGAGAACTGCTGCT	2120
Qy	1813	GAGGCGAGAGGAGTCTACGCCAGCAGGTGGAGGCTTGTGAGGGGGTGGCGCTTTCAGCC	1872
Db	2121	GAGGCGAGAGGAGTCTACGCCAGCAGGTGGAGGCTTGTGAGGGGGTGGCGCTTTCAGCC	2180
Qy	1873	CTCTGGCTTGGCCTTTGCTTCAACAGTGTAAATATCCCTCCCACTTCTCTTCCCTC	1932
Db	2181	CTCTGGCTTGGCCTTTGCTTCAACAGTGTAAATATCCCTCCCACTTCTCTTCCCTC	2240
Qy	1933	TCTTCCCTTCTCTCTCCCTCGGTGAATGATGGCTGCTTCTTAAACAAATACAACCA	1992
Db	2241	TCTTCCCTTCTCTCTCCCTCGGTGAATGATGGCTGCTTCTTAAACAAATACAACCA	2300
Qy	1993	AAACTCAGCAGTGTGATCTATAGCAGGATGGCCAGTACCTGGCTCCACTGATCACCTCT	2052
Db	2301	AAACTCAGCAGTGTGATCTATAGCAGGATGGCCAGTACCTGGCTCCACTGATCACCTCT	2360
Qy	2053	CTCTGTGACCAATCACCAACGGGTGCTTGTGGCTTGTGGCTTGTGGCTTGTGGCTTGTGGCT	2112
Db	2361	CTCTGTGACCAATCACCAACGGGTGCTTGTGGCTTGTGGCTTGTGGCTTGTGGCTTGTGGCT	2420
Qy	2113	TTCACTTGTGATCTGGGCTTCTTCTGTCATGCTGTAAGCTGTGGACGAGACCTTGA	2172
Db	2421	TTCACTTGTGATCTGGGCTTCTTCTTGTGTCATGCTGTAAGCTGTGGACGAGACCTTGA	2480
Qy	2173	CTTTTGTCTCTTAAAGGAAATGAGGAACTAAAGAAGTGAAGGGGTGGAGGGTGTGATC	2232
Db	2481	CTTTTGTCTCTTAAAGGAAATGAGGAACTAAAGAAGTGAAGGGGTGGAGGGTGTGATC	2540
Qy	2233	AGGGCAGTGTGAGCAGGGAGACCTCAGAGAAAGGCTGGAAAGGTGATTTCCCGTGTGA	2292
Db	2541	AGGGCAGTGTGAGCAGGGAGACCTCAGAGAAAGGCTGGAAAGGTGATTTCCCGTGTGA	2600

Qy 973 GCTGCTACTCTGGCGCAGTGCACTGTCAGCGTCTCCGTCTCTGTCCTGGGCTATGG 1032
 Db 1281 CTTGCTACTCTGGCGCAGTGCACTGTCAGCGTCTCCGTCTCTGTCCTGGGCTATGG 1340
 Qy 1033 GAAGTCCCTCCCTGGGCGAGCGTTTCGAGCAGGGGTCTTAGGCTCCCTGGCACTGGCAGG 1092
 Db 1341 GAAGTCCCTCCCTGGGCGAGCGTTTCGAGCAGGGGTCTTAGGCTCCCTGGCACTGGCAGG 1400
 Qy 1093 GCTGGCGCGGCACTGCGCCCTGGCTCAGTGGGAGAAATACGGGGCTCCCACTCTGGCT 1152
 Db 1401 GCTGGCGCGGCACTGCGCCCTGGCTCAGTGGGAGAAATACGGGGCTCCCACTCTGGCT 1460
 Qy 1153 GCCCTAGCGCGCACTGAGGGTTCAGCAGCAGCGCTGGGCTTCACCGTGGCCCTGGTGAT 1212
 Db 1461 GCCCTAGCGCGCACTGAGGGTTCAGCAGCAGCGCTGGGCTTCACCGTGGCCCTGGTGAT 1520
 Qy 1213 GATGAACCTCTCTGTTTCTGCTGGTGGTGGCGGCTCCCTACATCAAACTGTACTGTGACCT 1272
 Db 1521 GATGAACCTCTCTGTTTCTGCTGGTGGTGGCGGCTCCCTACATCAAACTGTACTGTGACCT 1580
 Qy 1273 GCCGGGGGCGACTTTGAGGCGGCTGTGGGACTGGCCATGTGTAGGCACTGGGCTGGCT 1332
 Db 1581 GCCGGGGGCGACTTTGAGGCGGCTGTGGGACTGGCCATGTGTAGGCACTGGGCTGGCT 1640
 Qy 1333 CATCTTCGACAGCGGCTCTCTACTGTCCTGGTGGCTTCTCTCAGCTTTCGGCTCCATGCT 1392
 Db 1641 CATCTTCGACAGCGGCTCTCTACTGTCCTGGTGGCTTCTCTCAGCTTTCGGCTCCATGCT 1700
 Qy 1393 GGGGCTCTTCCCTGTCAAGCCCGGAGGCGGTCAAGTCTGTCTCTGCTGGTGGTGGCTCCCT 1452
 Db 1701 GGGGCTCTTCCCTGTCAAGCCCGGAGGCGGTCAAGTCTGTCTCTGCTGGTGGTGGCTCCCT 1760
 Qy 1453 GCCTGCTGCTCAACCCACTGCTGTAGCTGCTCTTCAACCCCACTTCGGGATGACCT 1512
 Db 1761 GCCTGCTGCTCAACCCACTGCTGTAGCTGCTCTTCAACCCCACTTCGGGATGACCT 1820
 Qy 1513 TCGCGGCTTCGGCCCGCGCAGCGGACTCAGGCGCCCTAGCCTATGCTGGCGCGGCGGA 1572
 Db 1821 TCGCGGCTTCGGCCCGCGCAGCGGACTCAGGCGCCCTAGCCTATGCTGGCGCGGCGGA 1880
 Qy 1573 GCTGGAGAAGAGCTCTCTGTGATTTACCCAGGCGCTGGTAGCCTTCTCTGATGTGGATCT 1632
 Db 1881 GCTGGAGAAGAGCTCTCTGTGATTTACCCAGGCGCTGGTAGCCTTCTCTGATGTGGATCT 1940
 Qy 1633 CATCTGGAAGCTCTGAGCTGGGCGGCGCCCTGGGCTGGAGACCTATGGCTTCCCTC 1692
 Db 1941 CATCTGGAAGCTCTGAAAGCTGGGCGGCGCCCTGGGCTGGAGACCTATGGCTTCCCTC 2000
 Qy 1693 AGTGACCTCATCTCTGTCAGCAGCCAGGCGGCGCCCGAGGCTGGAGGCGCAGCCATTGCT 1752
 Db 2001 AGTGACCTCATCTCTGTCAGCAGCCAGGCGGCGCCCGAGGCTGGAGGCGCAGCCATTGCT 2060
 Qy 1753 AGAGCCAGAGGGGAACCACTTTGGGAACCCCAACCCCTCCATGGATGGAGAACTGCTGCT 1812
 Db 2061 AGAGCCAGAGGGGAACCACTTTGGGAACCCCAACCCCTCCATGGATGGAGAACTGCTGCT 2120
 Qy 1813 GAGGCGAGAGGATCTAGCCAGCAGGTGAGGCTTGTAGGGGGTGGCGGCTTTCAGCC 1872
 Db 2121 GAGGCGAGAGGATCTAGCCAGCAGGTGAGGCTTGTAGGGGGTGGCGGCTTTCAGCC 2180
 Qy 1873 CTCTGGCTGGGCTTTCCTTTCACACGCTGTAATAATCCCTCCCACTTCTCTTTCCTCCCTC 1932
 Db 2181 CTCTGGCTGGGCTTTCCTTTCACACGCTGTAATAATCCCTCCCACTTCTCTTTCCTCCCTC 2240
 Qy 1933 TCTTCCCTTTCCTCTCTCCCTCCGCTGAATGATGGCTGCTTCTTAAACAAATACAAACCA 1992
 Db 2241 TCTTCCCTTTCCTCTCTCCCTCCGCTGAATGATGGCTGCTTCTTAAACAAATACAAACCA 2300
 Qy 1993 AAACCTAGAGTGTGATCTATAGCAGGATGCCAGTACTGGCTCCACTGATCACTCT 2052
 Db 2301 AAACCTAGAGTGTGATCTATAGCAGGATGCCAGTACTGGCTCCACTGATCACTCT 2360
 Qy 2053 CTCCTGTGACCATCACCAACGGGTGCCTCTTGGGCTTTCCTTTCCTTTCCTTTCCTCAGC 2112

Db 2361 CTCTGTGACCATCACCAACGGGTGCCTCTTGGCTGGCTTTCCTTGGCTTCTCTCAGC 2420
 Qy 2113 TTCACTCTGATATCGGCGCTCTTCCCTTGTGTCATGTCGAACTGTGGACCGAGACCTGGA 2172
 Db 2421 TTCACTCTGATATCGGCGCTCTTCCCTTGTGTCATGTCGAACTGTGGACCGAGACCTGGA 2480
 Qy 2173 CTTTGTCTCTTAAAGGGAATGAGGAAGCTAAAGACAGTGAAGGCTGGAGGCTTGCATC 2232
 Db 2481 CTTTGTCTCTTAAAGGGAATGAGGAAGCTAAAGACAGTGAAGGCTGGAGGCTTGCATC 2540
 Qy 2233 AGGCGACAGTGGACAGGAGACCTCACAGAAAGGCTCGAAGGCTGATTTCCCTGTGA 2292
 Db 2541 AGGCGACAGTGGACAGGAGACCTCACAGAAAGGCTCGAAGGCTGATTTCCCTGTGA 2600
 Qy 2293 CTATGATAGGATACAAAATGTGTTCCATGTACCAATTAATCTTGAATGCAATGCAATGAT 2352
 Db 2601 CTATGATAGGATACAAAATGTGTTCCATGTACCAATTAATCTTGAATGCAATGCAATGAT 2660
 Qy 2353 AAAGACTTCTATTAATAAGCTTTGGAAGAGATTAAAAAAGGCTTAAAAA 2403
 Db 2661 AAAGACTTCTATTAATAAGCTTTGGAAGAGATTAAAAAAGGCTTAAAAA 2711

RESULT 9
 ABZ11841
 ID ABZ11841 standard; cDNA; 2988 BP.
 XX
 AC ABZ11841;
 XX
 DT 20-JAN-2003 (first entry)
 XX
 DE Human polynucleotide SEQ ID NO 723.
 XX
 KW Human; genome mapping; gene therapy; food supplement; virus; fungus;
 cell-proliferative disorder; neurodegenerative disease; bacterial;
 Parkinson's disease; Alzheimer's disease; autoimmune disease;
 multiple sclerosis; diabetes; genetic disorder; wound; burn; infection;
 arthritis; cytostatic; immunomodulator; nontropic; neuroprotective;
 antiparkinsonian; antidiabetic; immunosuppressive; dermatological;
 haemostatic; vulnery; fungicide; antibacterial; virucide; protozoacide;
 antiarthritic; gene; ss.
 XX
 OS Homo sapiens.
 XX
 PN WO200270539-A2.
 XX
 PD 12-SEP-2002.
 XX
 PF 05-MAR-2002; 2002WO-US005095.
 XX
 PR 05-MAR-2001; 2001US-00799451.
 XX
 PA (HYSE-) HYSEQ INC.
 XX
 PI Tang YT, Zhou P, Goodrich RW, Asundi V, Zhang J, Zhao QA, Ren F;
 PI Xue AJ, Yang Y, Yamazaki V, Chen R, Wang Z, Ghosh M;
 PI Wehrman T, Wang J, Wang D, Drmanac RT;
 XX
 DR WPI; 2002-759812/82.
 DR P-PSDB; ABP69624.
 XX
 PT New polynucleotides comprising sequences assembled from expressed
 sequence tags (ESTs), useful for treating cell-proliferative,
 neurodegenerative, autoimmune, genetic, myeloid or lymphoid, or platelet
 or coagulation disorders.
 XX
 PS Claim 1; SEQ ID NO 723; 1012pp + Sequence Listing; English.
 XX
 CC The invention relates to an isolated polynucleotide (I) comprising a
 nucleotide sequence selected from any of 948 sequences (ABZ1119-
 CC ABZ12066) or their mature protein coding portion, active domain coding
 protein or complementary sequences. The polynucleotides are useful for

CC identifying expressed genes or for physical mapping of human genome. The
 CC encoded polypeptides (ABP69802-ABP69849) are useful as molecular weight
 CC markers, as a food supplement, for generating antibodies, in medical
 CC imaging, screening and diagnostic assays and for treating cell-
 CC proliferative disorders (cancer), neurodegenerative diseases (Parkinson's
 CC or Alzheimer's disease), autoimmune diseases (multiple sclerosis,
 CC diabetes, lupus) genetic disorders, myeloid or lymphoid disorders,
 CC platelet or coagulation disorders, wound, burns, incision, ulcers, liver
 CC or lung fibrosis, infections (bacterial, viral, fungal, parasitic),
 CC arthritis, etc. Note: The sequence data for this patent did not form part
 CC of the printed specification, but was obtained in electronic format
 CC directly from WIPO at ftp.wipo.int/pub/published_pct_sequences
 XX
 SQ Sequence 2988 BP; 585 A; 924 C; 825 G; 654 T; 0 U; 0 Other;

Query Match 90.7%; Score 2255.8; DB 6; Length 2988;
 Best Local Similarity 99.9%; Pred. No. 0;
 Matches 2268; Conservative 0; Mismatches 2; Indels 1; Gaps 1;

QY 133 AGCGGCTCCAAACACAAACCGCATCTGGGAAATTTGAGCTNGACACTTACGCCAGCTGAG 192
 DB 718 ATCGGCTCCAAACACAAACCGCATCTGGGAAATTTGAGCT -GACACTTTCAGCCAGCTGAG 776
 QY 193 CTCCTGCAAGCCCTGGATCTTAGCTGGAACGCCATCCGCTCCATCCACCCTGAGGCTT 252
 DB 777 CTCCTGCAAGCCCTGGATCTTAGCTGGAACGCCATCCGCTCCATCCACCCTGAGGCTT 836
 QY 253 CTCACCTCGACTCCCTGGTCAAGCTGAGCTGAGCAACACAGCTGAGCCACACTGCC 312
 DB 837 CTCACCTCGACTCCCTGGTCAAGCTGAGCTGAGCAACACAGCTGAGCCACACTGCC 896
 QY 313 CTGCTGCACTTGGGGCTTGATGATCTGAAGCTCAAAGGAACTTGTCTCTCCCA 372
 DB 897 CTGCTGCACTTGGGGCTTGATGATCTGAAGCTCAAAGGAACTTGTCTCTCCCA 956
 QY 373 GGCCTTCTCAAGGACAGTTTCCAAACTGAGGATCTTGGAGTGCTTATGCCATCA 432
 DB 957 GGCCTTCTCAAGGACAGTTTCCAAACTGAGGATCTTGGAGTGCTTATGCCATCA 1016
 QY 433 GTGCTGCTCCCTATGGGATGTGCGAGCTTCTTCAAGGCTCTGCGGAGTGAGGCTGA 492
 DB 1017 GTGCTGCTCCCTATGGGATGTGCGAGCTTCTTCAAGGCTCTGCGGAGTGAGGCTGA 1076
 QY 493 AGACCTTCACTTGATGAGGAGTCTTCAAAAAGGCGCTTGGGCTCTTGGCAGACA 552
 DB 1077 AGACCTTCACTTGATGAGGAGTCTTCAAAAAGGCGCTTGGGCTCTTGGCAGACA 1136
 QY 553 AGCAGAGAACCATATGACAGGACCTGGATGAGCTCCAGCTGGAGATGGAGACTCAA 612
 DB 1137 AGCAGAGAACCATATGACAGGACCTGGATGAGCTCCAGCTGGAGATGGAGACTCAA 1196
 QY 613 GCCACACCCAGTGTCAGTGTCAGTCTTCCAGGCGCTTCAAGCCCTGTGAGTACCT 672
 DB 1197 GCCACACCCAGTGTCAGTGTCAGTCTTCCAGGCGCTTCAAGCCCTGTGAGTACCT 1256
 QY 673 CTTTGAAGCTGGGCACTCCGCTGGCGGTGCGGCTTGGGCTTGGCTCTCGCTGCTG 732
 DB 1257 CTTTGAAGCTGGGCACTCCGCTGGCGGTGCGGCTTGGGCTTGGCTCTCGCTGCTG 1316
 QY 733 CAATGGACTGTGCTGCTGACCGTGTTCGCTGGCGGCTGCGCCCTTGGCCCCGGGTCAA 792
 DB 1317 CAATGGACTGTGCTGCTGACCGTGTTCGCTGGCGGCTGCTGCGCCCTGCCCCGGGTCAA 1376
 QY 793 GTTTGTGTAGTGGGATGGAGGCGCAACACTTGTAGCTGGCATTTCTGTGGCTTCT 852
 DB 1377 GTTTGTGTAGTGGGATGGAGGCGGCAACACTTGTAGCTGGCATTTCTGTGGCTTCT 1436
 QY 853 AGCCTCAGTCAGTGCCTGACCTTGGTTCAGTTCCTGAGTACGAGGCGCGCTGGAGAC 912
 DB 1437 AGCCTCAGTCAGTGCCTGACCTTGGTTCAGTTCCTGAGTACGAGGCGCGCTGGAGAC 1496
 QY 913 GGGGCTAGGCTGCGGGCCACTGGGTTCTGCGAGTACTTGGGTCGAGGATCGGTGCT 972

DB 1497 GGGGCTAGGCTGCGGGGCCACTGGCTTCTTGGCAGTACTTGGGTCGAGGATCGGTGCT 1556
 QY 973 GCTGCTCACTCTGGGCGCAGTGCACTGAGGCTCTCCGCTCTCTCTGTCGGGCTATGG 1032
 DB 1557 GCTGCTCACTCTGGGCGCAGTGCACTGAGGCTCTCCGCTCTCTCTGTCGGGCTATGG 1616
 QY 1033 GAACTCCCTCTCTGGGCGCAGGTTTCGAGCAGGGGTCTTAGGCTCCCTGGCACTGGCAGG 1092
 DB 1617 GAACTCCCTCTCTGGGCGCAGGTTTCGAGCAGGGGTCTTAGGCTCCCTGGCACTGGCAGG 1676
 QY 1093 GCTGGCGCGCAGTCTGGGCTGAGTGGGAGAAACGGGGCTCTCCCACTCTGCT 1152
 DB 1677 GCTGGCGCGCAGTCTGGGCTGAGTGGGAGAAACGGGGCTCTCCCACTCTGCT 1736
 QY 1153 GCCCTACGCGCACCTGAGGCTCAGCAGCAGCCCTGGGCTTCAAGCTGGGCTGCTGAT 1212
 DB 1737 GCCCTACGCGCACCTGAGGCTCAGCAGCAGCCCTGGGCTTCAAGCTGGGCTGCTGAT 1796
 QY 1213 GATGAATCTCTCTGTTTCTGTTCTGTTGCGGCTTACATCAAACTGTACTGTGACT 1272
 DB 1797 GATGAATCTCTCTGTTTCTGTTCTGTTGCGGCTTACATCAAACTGTACTGTGACT 1856
 QY 1273 GCGCGGGGCGACTTTGAGGCGCTGTTGGGACTGTTGGGCTGTTGGGCTGTTGGCT 1332
 DB 1857 GCGCGGGGCGACTTTGAGGCGCTGTTGGGACTGTTGGGCTGTTGGGCTGTTGGCT 1916
 QY 1333 CATCTTGCAGACGGGCTCTCTGTTGCGGCTTCTCTGAGTTCGCTTCCATGCT 1392
 DB 1917 CATCTTGCAGACGGGCTCTCTGTTGCGGCTTCTCTGAGTTCGCTTCCATGCT 1976
 QY 1393 GGGCTCTTCTCTGTTGTTGCGGCTTCAAGTCTGTTCTGCTGTTGTTGTTGCTGCT 1452
 DB 1977 GGGCTCTTCTCTGTTGTTGCGGCTTCAAGTCTGTTCTGCTGTTGTTGTTGCTGCT 2036
 QY 1453 GCTGCTGCTCTCAACCCACTGTTGTTGCTTCTTCAACCCCACTTCCGGGATGACT 1512
 DB 2037 GCTGCTGCTCTCAACCCACTGTTGTTGCTTCTTCAACCCCACTTCCGGGATGACT 2096
 QY 1513 TCGCGGCTTTCGGGCGCGCAGGAGCTCAGGCGGCTTCTGCTGCTGCTGCTGCTGCT 1572
 DB 2097 TCGCGGCTTTCGGGCGCGCAGGAGCTCAGGCGGCTTCTGCTGCTGCTGCTGCTGCT 2156
 QY 1573 GCTGGAAGAGCTCTCTGTTGTTTACCCAGGCTTCTGTTGCTTCTGTTGTTGCT 1632
 DB 2157 GCTGGAAGAGCTCTCTGTTGTTTACCCAGGCTTCTGTTGCTTCTGTTGTTGCT 2216
 QY 1633 CATCTGGAAGCTCTGTTGAGCTGGGCGGCTTGGGCTGTTGAGCTTATGGCTTCCCTC 1692
 DB 2217 CATCTGGAAGCTCTGTTGAGCTGGGCGGCTTGGGCTGTTGAGCTTATGGCTTCCCTC 2276
 QY 1693 AGTGACCTCATCTCTGTTGAGCAGCAGGCGGCTTCCAGGCTTGGAGGCGCATTGCT 1752
 DB 2277 AGTGACCTCATCTCTGTTGAGCAGCAGGCGGCTTCCAGGCTTGGAGGCGCATTGCT 2336
 QY 1753 AGAGCCAGAGGAGAACCACTTTTGGGAAACCCCAACCTTCAATGGATGGAGAACTGCTGCT 1812
 DB 2337 AGAGCCAGAGGAGAACCACTTTTGGGAAACCCCAACCTTCAATGGATGGAGAACTGCTGCT 2396
 QY 1813 GAGGCGAGAGGATCTACGCGCAGAGTGGAGGCTTGTTCAGGGGTGCGGCTTTCAGCC 1872
 DB 2397 GAGGCGAGAGGATCTACGCGCAGAGTGGAGGCTTGTTCAGGGGTGCGGCTTTCAGCC 2456
 QY 1873 CTCTGGCTTGGCTTCTTCAACAGTGAATATCCCTCCCACTTCTCTCTTCTTCTTCT 1932
 DB 2457 CTCTGGCTTGGCTTCTTCAACAGTGAATATCCCTCCCACTTCTCTCTTCTTCTTCT 2516
 QY 1933 TCTTCT 1992
 DB 2517 TCTTCT 2576
 QY 1993 AAATCTAGCAGTGTGATCTATAGCAGATGCGCCAGTACTGTTGCTTCACTGATCACTCT 2052
 DB 2577 AAATCTAGCAGTGTGATCTATAGCAGATGCGCCAGTACTGTTGCTTCACTGATCACTCT 2636

QY 2053 CTCCTGTGACCATCAACACGGGTGCTCTTTGGCCCTGGCTTTCCCTTGGCCTTCTCTCAGC 2112
 DB |||||
 QY 2637 CTCCTGTGACCATCAACACGGGTGCTCTTTGGCCCTGGCTTTCCCTTGGCCTTCTCTCAGC 2696
 DB |||||
 QY 2113 TTCACCTTGATACCTGGGCTCTTCTCTGTCATGCTGAAAGCTGTGGACGAGACCTGGA 2172
 DB |||||
 QY 2697 TTCACCTTGATACCTGGGCTCTTCTCTGTCATGCTGAAAGCTGTGGACGAGACCTGGA 2756
 DB |||||
 QY 2173 CTTTGTCTGCTTAAGGGAATGAGGAAGTAAGACAGATGAGGGGTGGAGGTTGATC 2232
 DB |||||
 QY 2233 AGGGCACAGTGGACAGGGAGACCTCACAGAGAAAGGCTGGAAGTGATTTCCCGTGTGA 2292
 DB |||||
 QY 2817 AGGGCACAGTGGACAGGGAGACCTCACAGAGAAAGGCTGGAAGTGATTTCCCGTGTGA 2876
 DB |||||
 QY 2293 CTCATGGATAGGATACAAATGTGTTCATGTACATTAATCTTTGACATATGCCATGCAT 2352
 DB |||||
 QY 2877 CTCATGGATAGGATACAAATGTGTTCATGTACATTAATCTTTGACATATGCCATGCAT 2936
 DB |||||
 QY 2353 AAAGACTTCTTAAATAAGCTTTGGAAGAGATTAAAAA 2403
 DB |||||
 QY 2937 AAAGACTTCTTAAATAAGCTTTGGAAGAGATTAAAAA 2987
 DB |||||

RESULT 10

ADM44359

ID ADM44359 standard; cDNA; 2988 BP.

XX AC ADM44359;

XX DT 03-JUN-2004 (first entry)

DE Novel human arginine-rich protein cDNA #723.

XX ss; Gene; human; arginine-rich protein; cancer; inflammation;

XX genetic disorder.

XX OS Homo sapiens.

XX US2004053250-A1.

XX 18-MAR-2004.

XX 21-NOV-2002; 2002US-00302172.

XX 05-MAR-2001; 2001US-00799451.

XX 05-MAR-2002; 2002WO-05005095.

XX 20-AUG-2002; 2002US-00225251.

XX (TANG/) TANG Y T.

PA (XUEA/) XUE A.

PA (DRMA/) DRMANAC R T.

XX Tang YT, Xue A, Drmanac RT;

XX WPI; 2004-238579/22.

XX New isolated arginine-rich protein-like polynucleotides and polypeptides,

XX useful for diagnosing and/or treating conditions associated with aberrant

XX activity of the arginine-rich polypeptides, such as cancer and

XX inflammation.

XX Disclosure; SEQ ID NO 723; 51pp; English.

XX The invention relates to an isolated polynucleotide. The methods and

XX compositions of the present invention are useful for the diagnosis and/or

XX treatment of diseases or conditions associated with aberrant expression

XX or activity of the arginine-rich protein-like polypeptides, such as

XX cancer and inflammation. They can also be used in forensics, gene

XX mapping, identification of mutations responsible for genetic disorders,

XX and in assessing biodiversity. The present sequence represents a novel

CC human arginine-rich protein cDNA.

XX Sequence 2988 BP; 585 A; 924 C; 825 G; 654 T; 0 U; 0 Other;

SQ Query Match 90.7%; Score 2255.8; DB 12; Length 2988;
 Best Local Similarity 99.9%; Pred. No. 0;
 Matches 2268; Conservative 0; Mismatches 2; Indels 1; Gaps 1;

QY 133 AGCGGCTCCAAACACACCGCATCTGGGAAATTTGGAGCTTGACACCTTCAGCCAGCTGAG 192
 DB |||||
 QY 718 ATCGGCTCCAAACACACCGCATCTGGGAAATTTGGAGCT-GACACCTTCAGCCAGCTGAG 776
 DB |||||
 QY 193 CTCCTCTCAAGCCCTGGATCTTAGCTGGAACCGCATCCGGTCCATCCACCCCTGAGGCTT 252
 DB |||||
 QY 777 CTCCTCTCAAGCCCTGGATCTTAGCTGGAACCGCATCCGGTCCATCCACCCCTGAGGCTT 836
 DB |||||
 QY 253 CTCACCCCTGACATCTCCCTGTCAAGCTTGAGCTTGACAGACCAACCAAGCTGACCACTGCC 312
 DB |||||
 QY 837 CTCACCCCTGACATCTCCCTGTCAAGCTTGAGCTTGAGAGCTCAAAAGGAACTTGTCTCTCCCA 896
 DB |||||
 QY 313 CTTGGCTTGAATTTGGGGCTTGATGATCTGAGCTCAAAAGGAACTTGTCTCTCTCCCA 372
 DB |||||
 QY 897 CTTGGCTTGAATTTGGGGCTTGATGATCTGAGCTCAAAAGGAACTTGTCTCTCTCCCA 956
 DB |||||
 QY 373 GGCCTTCTCAAGGACAGATTTCCCAAACTGAGGATCTTGAGGTGCTTTATGCTTACCA 432
 DB |||||
 QY 957 GGCCTTCTCAAGGACAGATTTCCCAAACTGAGGATCTTGAGGTGCTTTATGCTTACCA 1016
 DB |||||
 QY 433 GTGCTGTCCATATGGGATGTGCCAGCTTCTTCAAGGCTCTTGGGAGTGGGAGGCTGA 492
 DB |||||
 QY 1017 GTGCTGTCCATATGGGATGTGCCAGCTTCTTCAAGGCTCTTGGGAGTGGGAGGCTGA 1076
 DB |||||
 QY 493 AGACCTTCACTTTGATGATGAGGATCTTCAAAAGGCTCTTGGGCTCTCTTGGCAGACA 552
 DB |||||
 QY 1077 AGACCTTCACTTTGATGATGAGGATCTTCAAAAGGCTCTTGGGCTCTCTTGGCAGACA 1136
 DB |||||
 QY 553 AGACGAGAACACCTATGACACGACCTTGGATGAGCTCCAGCTGGAGATGGAGGACTCAA 612
 DB |||||
 QY 1137 AGACGAGAACACCTATGACACGACCTTGGATGAGCTCCAGCTGGAGATGGAGGACTCAA 1196
 DB |||||
 QY 613 GCACACCCAGTGTCCAGTGTAGCCCTACTCCAGGCCCTTCAAGCCCTGTGAGTACCT 672
 DB |||||
 QY 1197 GCACACCCAGTGTCCAGTGTAGCCCTACTCCAGGCCCTTCAAGCCCTGTGAGTACCT 1256
 DB |||||
 QY 673 CTTTGAAGCTGGGGCATCCGCTGGCCGTGCGGCTCATCGTGTGCTCTCCGCTCTGTG 732
 DB |||||
 QY 1257 CTTTGAAGCTGGGGCATCCGCTGGCCGTGCGGCTCATCGTGTGCTCTCCGCTCTGTG 1316
 DB |||||
 QY 733 CAATGGACTGGTCTGCTGACCGTGTTCGCTGGCGGCTTGGCCCTTGGCCCTGGCCGCTCAA 792
 DB |||||
 QY 1317 CAATGGACTGGTCTGCTGACCGTGTTCGCTGGCGGCTTGGCCCTTGGCCCTGGCCGCTCAA 1376
 DB |||||
 QY 793 GTTTGTGTAGGTGCGATTCGAGCGCCAAACACTTGTGACTGGCAATTTCTGTGGCTTCT 852
 DB |||||
 QY 1377 GTTTGTGTAGGTGCGATTCGAGCGCCAAACACTTGTGACTGGCAATTTCTGTGGCTTCT 1436
 DB |||||
 QY 853 AGCCTCAGTCGATCCCTGACCTTTGGTCACTTCTCTGAGTACGGAGCCGCTGGGAGAC 912
 DB |||||
 QY 1437 AGCCTCAGTCGATCCCTGACCTTTGGTCACTTCTCTGAGTACGGAGCCGCTGGGAGAC 1496
 DB |||||
 QY 913 GGGCTAGGCTGCGGGGCCACTCGCTTCCCTGGCAGTACTTTGGGTGGAGGACATCGGTGCT 972
 DB |||||
 QY 1497 GGGCTAGGCTGCGGGGCCACTCGCTTCCCTGGCAGTACTTTGGGTGGAGGACATCGGTGCT 1556
 DB |||||
 QY 973 GCTGCTCACTTCTGGCCGCACTGAGTGCAGTGCAGCGTCTCCGTCTCTGTGTCGGGCTATGG 1032
 DB |||||
 QY 1557 GCTGCTCACTTCTGGCCGCACTGAGTGCAGTGCAGCGTCTCCGTCTCTGTGTCGGGCTATGG 1616
 DB |||||
 QY 1033 GAAGTCCCTTCTGGCCGCACTTGCAGTGCAGGCTTCTAGGCTGCTGCTGCACTGGCAGG 1092
 DB |||||
 QY 1617 GAAGTCCCTTCTGGCCGCACTTGCAGTGCAGGCTTCTAGGCTGCTGCTGCACTGGCAGG 1676
 DB |||||
 QY 1093 GCTGGCGCGGCACTGCCCCCTGGCCCTCAGTGGGAGAAATACGGGGCTTCCCCACTCTGCCT 1152
 DB |||||

used as an antagonist of the LGR6 signalling pathway. The invention provides LGR6.1 and LGR6.2 polypeptides and nucleic acids, as well as selective binding agents, vectors, host cells and methods for producing LGR6-SV polypeptides. Also provided are pharmaceutical compositions and methods for the diagnosis, treatment, amelioration and/or prevention of diseases, disorders and conditions associated with LGR6-SV polypeptides. LGR6-SV polypeptides and nucleic acids can be used for diagnosing or treating: cancers such as hepatocellular carcinoma, adenomatous polypsis, malignant transformation of epithelial cells, colorectal cancer, breast cancer, ovarian cancer, etc.; cachexia and muscular dystrophy; miscarriage, endometriosis, uterine cancer, and female infertility; Cushing's disease and Addison's disease; male infertility and testicular cancer; leukaemia; and diseases and conditions affecting the kidney. They may also be useful for treating diseases and conditions involving cell proliferation and differentiation, such as tissue damage and degeneration, ageing, wound healing, cancer, hyperplasia and hypertrophy.

XX SQ Sequence 3306 BP; 669 A; 1031 C; 885 G; 721 T; 0 U; 0 Other;

Query Match		90.4%;	Score 2247.8;	DB 12;	Length 3306;
Best Local Similarity		99.6%;	Pred. No. 0;		
Matches 2263;		Conservative 0;	Mismatches 7;	Indels 1;	Gaps 1;
Qy	133	AGCGGCTCCAAACACACCGCATCTGGGAATTTGGAGCTNGACACCTTCAGCCAGCTGAG	192		
Db	1025	ATCGGCTTCCAAACACACCGCATCTGGGAATTTGGAGCT-GACACCTTCAGCCAGCTGAG	1083		
Qy	193	CTCCCTGCAAGCCCTGGATCTTAGCTGGAACGCCATCCGGTCCATCCACCTCAGGCGCTT	252		
Db	1084	CTCCCTGCAAGCCCTGGATCTTAGCTGGAACGCCATCCGGTCCATCCACCCCGAGGCTT	1143		
Qy	253	CTCCACCTGCACTCCCTGCTCAAGCTGGAGCTGACAGAACACAGCTGACCACTGCC	312		
Db	1144	CTCCACCTGCACTCCCTGCTCAAGCTGGAGCTGACAGAACACAGCTGACCACTGCC	1203		
Qy	313	CCTGCTGGACTTGGGGCTTGATGCATCTCAAGCTCAAGGGAACTTGCTCTCCCA	372		
Db	1204	CCTGCTGGACTTGGGGCTTGATGCATCTCAAGCTCAAGGGAACTTGCTCTCCCA	1263		
Qy	373	GGCCTTCTCCAAAGACAGTTCCTCCAAACCTGAGGATCCTGGAGTGCTTATGCTACCA	432		
Db	1264	GGCCTTCTCCAAAGACAGTTCCTCCAAACCTGAGGATCCTGGAGTGCTTATGCTACCA	1323		
Qy	433	GTGCTGCCCTATGGGATGTGTCAGCTTCTTCAAGGCTCTGGGAGTGGGAGGCTGA	492		
Db	1324	GTGCTGCCCTATGGGATGTGTCAGCTTCTTCAAGGCTCTGGGAGTGGGAGGCTGA	1383		
Qy	493	AGACCTTCACCTTGATGAGGAGTCTTCAAAAGGCGCCCTGGGCTCCTTGGCCAGACA	552		
Db	1384	AGACCTTCACCTTGATGAGGAGTCTTCAAAAGGCGCCCTGGGCTCCTTGGCCAGACA	1443		
Qy	553	AGCAGAGAACCACATATGACACAGGACCTGGATGAGCTCCAGCTGGAGATGGAGGACTCAAA	612		
Db	1444	AGCAGAGAACCACATATGACACAGGACCTGGATGAGCTCCAGCTGGAGATGGAGGACTCAAA	1503		
Qy	613	GCACACCCAGTGTCCAGTGTAGCCCTATCTCCAGGCGCCCTTCAAGCCCTGTGAGTACCT	672		
Db	1504	GCACACCCAGTGTCCAGTGTAGCCCTATCTCCAGGCGCCCTTCAAGCCCTGTGAGTACCT	1563		
Qy	673	CTTTGAAAGCTGGGGCATCGCCCTGGCGGTGGGGCCATCGTGTGCTCCCGTCTG	732		
Db	1564	CTTTGAAAGCTGGGGCATCCGCCCTGGCGGTGGGGCCATCGTGTGCTCCCGTCTG	1623		
Qy	733	CAATGGAGTGGTGTGCTGACCGGTGTCGCTGGCGGGCTGCCCCCTGCCCGCCGGTCAA	792		
Db	1624	CAATGGAGTGGTGTGCTGACCGGTGTCGCTGGCGGGCTGTGCCCTGCCCGCCGGTCAA	1683		
Qy	793	GTTTGTGGTAGTTCGGATGGAGCGGCAACACCTTGACTGGGATTTCTGTGGCCCTTCT	852		
Db	1684	GTTTGTGGTAGTTCGGATGGAGCGGCAACACCTTGACTGGGATTTCTGTGGCCCTTCT	1743		
Qy	853	AGCCTCAGTCGATGCCCTTGACCTTTTGTGTCAGTTTCTCTGAGTGACGAGCGCCGCTGGAGAC	912		

Db	1744	AGCCTCAGTCGATGCCCTTGACCTTTTGTGTCAGTTTCTCTGAGTACGAGAGCCCTGGAGAC	1803		
Qy	913	GGGGCTAGGTGTCGGGGCCACTGGCTTCTTGCGCAGTACTTGGGTGCGAGGCACTGGGTGCT	972		
Db	1804	GGGGCTAGGTGTCGGGGCCACTGGCTTCTTGCGCAGTACTTGGGTGCGAGGCACTGGGTGCT	1863		
Qy	973	GCTGCTCACTCTGGCCGAGTGCAGTGCAGGCTTCCGTCCTCCTCTCTGTCGCCGGGCTATGG	1032		
Db	1864	GCTGCTCACTCTGGCCGAGTGCAGTGCAGGCTTCCGTCCTCCTCTCTGTCGCCGGGCTATGG	1923		
Qy	1033	GAAGTCCCTCCCTGTCGGCAGCGTTCGAGCAGGGGTCTTAGGCTGCCCTGGCAGCTGGCAGG	1092		
Db	1924	GAAGTCCCTCCCTGTCGGCAGCGTTCGAGCAGGGGTCTTAGGCTGCCCTGGCAGCTGGCAGG	1983		
Qy	1093	GCTGGCCGCGCAGCTGCCCTTGCGCTCAGTGGGAGAAATACGGGGCCTTCCCACTCTGCT	1152		
Db	1984	GCTGGCCGCGCAGCTGCCCTTGCGCTCAGTGGGAGAAATACGGGGCCTTCCCACTCTGCT	2043		
Qy	1153	GCCTAGCGCCCACTGAGGTCAGCCAGAGCCCTGGGCTTCAACCTGGGCCCTGGTGAT	1212		
Db	2044	GCCTAGCGCCCACTGAGGTCAGCCAGAGCCCTGGGCTTCAACCTGGGCCCTGGTGAT	2103		
Qy	1213	GATGAACCTCTTCTGTTTCTGTCGTCGGCGGTGCTTACATCAAACTGTACTGTGACCT	1272		
Db	2104	GATGAACCTCTTCTGTTTCTGTCGTCGGCGGTGCTTACATCAAACTGTACTGTGACCT	2163		
Qy	1273	GCCGCGGGCGACTTTGAGGCCGTCGCGCATGCGCCATGGTGGAGCACTGGGCTCGGCT	1332		
Db	2164	GCCGCGGGCGACTTTGAGGCCGTCGCGCATGCGCCATGGTGGAGCACTGGGCTCGGCT	2223		
Qy	1333	CATCTTGCAGAGCGGCTCTACTGTCGCGTGGCCCTTCTCAGCTTCGCTCCCATGCT	1392		
Db	2224	CATCTTGCAGAGCGGCTCTCTACTGTCGCGTGGCCCTTCTCAGCTTCGCTCCCATGCT	2283		
Qy	1393	GGGCTCTTCTCCCTGTCACGCCCGAGGCGCTCAAGTCTGTCTGCTGTGTGTGCTGCCCT	1452		
Db	2284	GGGCTCTTCTCCCTGTCACGCCCGAGGCGCTCAAGTCTGTCTGCTGTGTGTGCTGCCCT	2343		
Qy	1453	GCTGCTGCTCTCAACCACTGTGTACTGTCTTCAACCCCACTTCCGCGGATGACCT	1512		
Db	2344	GCTGCTGCTCTCAACCACTGTGTACTGTCTTCAACCCCACTTCCGCGGATGACCT	2403		
Qy	1513	TCGGCGCTTTCGGCCCGCGCAGGGGACTCAGGGCCCTAGCCTATGCTCGGGCGGGGA	1572		
Db	2404	TCGGCGCTTTCGGCCCGCGCAGGGGACTCAGGGCCCTAGCCTATGCTCGGGCGGGGA	2463		
Qy	1573	GCTGGAGAGAGCTCTCTGTGATTTCTACCCAGGCGCTGGTAGCCTTCTCTGATGTGGATCT	1632		
Db	2464	GCTGGAGAGAGCTCTCTGTGATTTCTACCCAGGCGCTGGTAGCCTTCTCTGATGTGGATCT	2523		
Qy	1633	CATCTGGAAGCTTCTGAAGCTGGGCGGCCCTCTGGCTGGAGACCTATGGCTTCCCTC	1692		
Db	2524	CATCTGGAAGCTTCTGAAGCTGGGCGGCCCTCTGGCTGGAGACCTATGGCTTCCCTC	2583		
Qy	1693	AGTGACCTCATCTCTGTGTCAGCAGCCAGGGGCCCCCAGGCTGGAGGCGACCACTTCTGT	1752		
Db	2584	AGTGACCTCATCTCTGTGTCAGCAGCCAGGGGCCCCCAGGCTGGAGGCGACCACTTGTGT	2643		
Qy	1753	AGAGCCAGAGGGGAAACACCTTTTGGGAACCCCAACCTCCATGGATGGAGAACTGTGCT	1812		
Db	2644	AGAGCCAGAGGGGAAACACCTTTTGGGAACCCCAACCTCCATGGATGGAGAACTGTGCT	2703		
Qy	1813	GAGGCGAGAGGATCTACGCCAGCAGGTGAGGCTTGTACAGGGGTGGCGGCTTTCAGCC	1872		
Db	2704	GAGGCGAGAGGATCTACGCCAGCAGGTGAGGCTTGTACAGGGGTGGCGGCTTTCAGCC	2763		
Qy	1873	CTCTGGCTTGGCTTGTGCTTCAACGTTGTAATATCCCTCCCACTTCTCTCTCCCTC	1932		
Db	2764	CTCTGGCTTGGCTTGTGCTTCAACGTTGTAATATCCCTCCCACTTCTCTCTCCCTC	2823		
Qy	1933	TCCTTCCCTTCTCTCTCCCTCCTGGTGAATGATGGCTGCTTCTTAAACAAATAACAACA	1992		

Qy 673 CTTTGAAGCTGGGCATCCGCTGGCGTGTGGCCATCGTGTGCTCTCCGTGCTCTG 732
 Db 1599 CTTTGAAGCTGGGCATCCGCTGGCGTGTGGCCATCGTGTGCTCTCCGTGCTCTG 1658
 Qy 733 CAATGGACTGTGTCTGAACCGTGTTCGTGGCGGCTCTGCCCTCTGCCCCGGGTCAA 792
 Db 1659 CAATGGACTGTGTCTGAACCGTGTTCGTGGCGGCTCTGCCCTCTGCCCCGGGTCAA 1718
 Qy 793 GTTTGTGTAGGTGGATTGGCAGGGCCAAACACCTTGACTGGCATTTCTGTGGCTTCT 852
 Db 1719 GTTTGTGTAGGTGGATTGGCAGGGCCAAACACCTTGACTGGCATTTCTGTGGCTTCT 1778
 Qy 853 AGCCTCAGTCAGTCCCTGACCTTTGGTTCAGTTCCTCTGAGTACGGAGCCCGCTGGAGAC 912
 Db 1779 AGCCTCAGTCAGTCCCTGACCTTTGGTTCAGTTCCTCTGAGTACGGAGCCCGCTGGAGAC 1838
 Qy 913 GGGGTAGGCTGCCGGGCCACTGGGTTCCCTGGCAGTACTTGGGTTCGAGGCAATCGGTGCT 972
 Db 1839 GGGGTAGGCTGCCGGGCCACTGGGTTCCCTGGCAGTACTTGGGTTCGAGGCAATCGGTGCT 1898
 Qy 973 GCTGCTCACTCTGGCGCAGTGCAGTGCAGCGTCTCCGTCTCCTGTGTCCGGGCTATGG 1032
 Db 1899 GCTGCTCACTCTGGCGCAGTGCAGTGCAGCGTCTCCGTCTCCTGTGTCCGGGCTATGG 1958
 Qy 1033 GAAGTCCCGCTCCCTGGGCAGCGTTTCGAGCAGGGGTCTTAGGCTGCCCTGGCACTGGCAGG 1092
 Db 1959 GAAGTCCCGCTCCCTGGGCAGCGTTTCGAGCAGGGGTCTTAGGCTGCCCTGGCACTGGCAGG 2018
 Qy 1093 GCTGGCGCGCCACTGGCCCTGGGCTCAGTGGGAGAAATACGGGGCCCTCCCACTCTGCCT 1152
 Db 2019 GCTGGCGCGCGCTGGCCCTGGGCTCAGTGGGAGAAATACGGGGCCCTCCCACTCTGCCT 2078
 Qy 1153 GCCCTAGCGGCCACTGAGGGTTCAGCCAGCAGCCCTGGGCTTTCACCGTGGGCCCTGGTGAT 1212
 Db 2079 GCCCTAGCGGCCACTGAGGGTTCAGCCAGCAGCCCTGGGCTTTCACCGTGGGCCCTGGTGAT 2138
 Qy 1213 GATGAATCCTCTCTGTTTCTCGTCTGTCGGCGGCTGCCCTACATCAAACTGTACTGTGACCT 1272
 Db 2139 GATGAATCCTCTCTGTTTCTCGTCTGTCGGCGGCTGCCCTACATCAAACTGTACTGTGACCT 2198
 Qy 1273 GCCCGGGGGCACTTTGAGGCGGTGTGGGACTGCGGCCATGTGAGGCACTGTGGCTGGCT 1332
 Db 2199 GCCCGGGGGCACTTTGAGGCGGTGTGGGACTGCGGCCATGTGAGGCACTGTGGCTGGCT 2258
 Qy 1333 CATCTTGCAGACGGGCTCTCTACTGTCCCGTGGCCCTTCTCCTCAGCTTCGCTCCATGCT 1392
 Db 2259 CATCTTGCAGACGGGCTCTCTACTGTCCCGTGGCCCTTCTCCTCAGCTTCGCTCCATGCT 2318
 Qy 1393 GGGCTCTTCTCCTGTCAAGCGGCGGTCAAGTCTCTGCTGTGGTGGTGTGCGCCCT 1452
 Db 2319 GGGCTCTTCTCCTGTCAAGCGGCGGTCAAGTCTCTGCTGTGGTGGTGTGCGCGT 2378
 Qy 1453 GCCTGCTGCTCAACCCACTGTGTACTGCTCTTCAACCCCACTTCCGGGATGACCT 1512
 Db 2379 GCCTGCTGCTCAACCCACTGTGTACTGCTCTTCAACCCCACTTCCGGGATGACCT 2438
 Qy 1513 TCGGGGCTTTCGGGCGCGCAGGGGACTCAGGGCCCTAGCCTATGCTGGGCGCGGGGA 1572
 Db 2439 TCGGGGCTTTCGGGCGCGCAGGGGACTCAGGGCCCTAGCCTATGCTGGGCGCGGGGA 2498
 Qy 1573 GCTGGAGAAGAGCTCTCTGTGATTCTACCCAGGCGCTGTGAGCTTCTCTGATGTGGATCT 1632
 Db 2499 GCTGGAGAAGAGCTCTCTGTGATTCTACCCAGGCGCTGTGAGCTTCTCTGATGTGGATCT 2558
 Qy 1633 CATCTTGAAGCTTCTGAAGCTGGGCGGCCCTTGGGCTGGAGACCTATGGCTTCCCGCTC 1692
 Db 2559 CATCTTGAAGCTTCTGAAGCTGGGCGGCCCTTGGGCTGGAGACCTATGGCTTCCCGCTC 2618
 Qy 1693 AGTGACCTCTATCTCTGTCTGAGCAGCCAGGGGCCCCCAGGCTGGAGGGCAGCCATTGTGT 1752
 Db 2619 AGTGACCTCTATCTCTGTCTGAGCAGCCAGGGGCCCCCAGGCTGGAGGGCAGCCATTGTGT 2678
 Qy 1753 AGAGCCAGGGGAAACCACTTTGGGAACCCCAACCCCTCCATGGATGGAGAACTGCTGCT 1812

Db 2679 AGAGCCAGAGGGGAACCACTTTGGGAACCCCAACCCCTCCATGGATGGAGAACTGCTGCT 2738
 Qy 1813 GAGGGCAGAGGATCTACGCCAGCAGGTGGAGGCTTGTTCAGGGGGTGGCGGCTTTCAGCC 1872
 Db 2739 GAGGGCAGAGGATCTACGCCAGCAGGTGGAGGCTTGTTCAGGGGGTGGCGGCTTTCAGCC 2798
 Qy 1873 CTCTGGGCTTGGCCTTTGCTTTCACACGTGTAATAATATCCTCCCATTTCTCTCTTCCCTC 1932
 Db 2799 CTCTGGGCTTGGCCTTTGCTTTCACACGTGTAATAATATCCTCCCATTTCTCTCTTCCCTC 2858
 Qy 1933 TCTTCCCTTTCTCTCTCTCCCTCGGTGAATGATGGCTGCTTCTTAAACAATAACAACCA 1992
 Db 2859 TCTTCCCTTTCTCTCTCTCCCTCGGTGAATGATGGCTGCTTCTTAAACAATAACAACCA 2918
 Qy 1993 AAACCTCAGCAGTGTGATCTATAGCAGGATGGCCAGTACCTGGCTCCACTGATCACCTCT 2052
 Db 2919 AAACCTCAGCAGTGTGATCTATAGCAGGATGGCCAGTACCTGGCTCCACTGATCACCTCT 2978
 Qy 2053 CTCTGTGACCATCACCAACGGGTGCTCTTTGGGCTGGCTTTCCCTTTGGGCTTCTCTCAGC 2112
 Db 2979 CTCTGTGACCATCACCAACGGGTGCTCTTTGGGCTGGCTTTCCCTTTGGGCTTCTCTCAGC 3038
 Qy 2113 TTCACTTGTATCTGGGCTCTTCTTGTCTCATGTCTGAAGCTGTGGACCGAGACCTGGA 2172
 Db 3039 TTCACTTGTATCTGGGCTCTTCTTGTCTCATGTCTGAAGCTGTGGACCGAGACCTGGA 3098
 Qy 2173 CTTTGTCTGCTTAAAGGAAATAGGGAATAGGGAAGTAAAGACAGTGAAGGGTGGAGGTTGATC 2232
 Db 3099 CTTTGTCTGCTTAAAGGAAATAGGGAAGTAAAGACAGTGAAGGGTGGAGGTTGATC 3158
 Qy 2233 AGGCAAGTGGACAGGAGACCTTCACAGAGAAAGGCTCGAAGGTGATTTCCCGTGTGA 2292
 Db 3159 AGGCAAGTGGACAGGAGACCTTCACAGAGAAAGGCTCGAAGGTGATTTCCCGTGTGA 3218
 Qy 2293 CTATGATAGGATACAAAATGTTCATGTACCATTAATTTTGACATATGCCATGCAT 2352
 Db 3219 CTATGATAGGATACAAAATGTTCATGTACCATTAATTTTGACATATGCCATGCAT 3278
 Qy 2353 AAAGACTTCTATTAATAAAGCTTTGGAAGAGATTTAAAAAATAAAAAA 2403
 Db 3279 AAAGACTTCTATTAATAAAGCTTTGGAAGAGATTTAAAAAATAAAAAA 3329

RESULT 13
 ADD89048
 ID ADD89048 standard; cDNA; 3325 BP.
 XX
 AC ADD89048;
 XX
 DT 29-JAN-2004 (first entry)
 XX
 DE Encoding sequence TAT251.
 XX
 KW tumour-associated antigenic target polypeptide; Cytostatic; tumour;
 KW cancer; 88.
 XX
 OS Homo sapiens.
 XX
 PN WO2003057160-A2.
 XX
 PD 17-JUL-2003.
 XX
 PF 30-DEC-2002; 2002WO-US041798.
 XX
 PR 02-JAN-2002; 2002US-0345444P.
 PR 25-JAN-2002; 2002US-0351885P.
 PR 25-FEB-2002; 2002US-0360066P.
 PR 03-MAR-2002; 2002US-0362004P.
 PR 20-MAR-2002; 2002US-0366869P.
 PR 21-MAR-2002; 2002US-0366284P.
 PR 28-MAR-2002; 2002US-0368679P.
 PR 19-AUG-2002; 2002US-0404809P.

PR	21-AUG-2002; 2002US-0405645P.	
XX	PA (GETH) GENENTECH INC.	
XX	PI Prantz G, Hillan KJ, Phillips H, Polakis P, Smith V, Spencer SD;	
XX	PI Williams PM, Wu TD, Zhang Z;	
XX	WPI; 2003-569537/53.	
DR	P-PSDB; ADD89064.	
XX	New antibodies against tumor-associated antigenic target polypeptide,	
PT	useful for treating or diagnosing tumors or cancers in mammals, e.g.	
PT	prostate cancer, lung cancer, prostate adenocarcinomas or renal cell	
PT	carcinomas.	
XX	Claim 1; SEQ ID NO 52; 252pp; English.	
XX	The present invention relates to antibodies against tumour-associated	
CC	antigenic target polypeptide. The antibody is useful for treating or	
CC	diagnosing tumours or cancers in mammals, e.g. prostate cancer, lung	
CC	cancer, breast cancer, colon cancer, ovarian cancer, prostate	
CC	adenocarcinomas, renal cell carcinomas, or pleural mesothelioma. The	
CC	present sequence represents a TAT polypeptide encoding sequence.	
XX	Sequence 3325 BP; 663 A; 1042 C; 893 G; 727 T; 0 U; 0 Other;	
SQ		
	Query Match 90.4%; Score 2247.2; DB 10; Length 3325;	
	Best Local Similarity 99.8%; Pred. No. 0;	
	Matches 2260; Conservative 0; Mismatches 3; Indels 1; Gaps 1;	
Qy	133 AGCGGCTCCACACACCGGATCTGGGAAATTGGAGCTNAGACCTTCAGCCAGCTGAG	192
Db	1063 ATCGGCTTCCACACACCGGATCTGGGAAATTGGAGCT-GACACCTTCAGCCAGCTGAG	1121
Qy	193 CTCCTGCAAGCCCTGGATCTTAGCTGGAACGCCATCCGCTCCATCCACCTGAGGCTT	252
Db	1122 CTCCTGCAAGCCCTGGATCTTAGCTGGAACGCCATCCGCTCCATCCACCTGAGGCTT	1181
Qy	253 CTCACCCCTGCACCTCCCTGGTCAAGCTGGACCTGCACAGAACACCACTGACACACTGCC	312
Db	1182 CTCACCCCTGCACCTCCCTGGTCAAGCTGGACCTGCACAGAACACCACTGACACACTGCC	1241
Qy	313 CTGGCTGGACTTGGGGCTTGATGCACTCTGAAGCTCAAGGAAACCTTGCTCTCTCCCA	372
Db	1242 CTGGCTGGACTTGGGGCTTGATGCACTCTGAAGCTCAAGGAAACCTTGCTCTCTCCCA	1301
Qy	373 GGCTTCTCCAGGACAGTTTCCAAACTGAGGATCTGGAGTGCGCTTATGCTATGCTACCA	432
Db	1302 GGCTTCTCCAGGACAGTTTCCCAAACTGAGGATCTGGAGTGCGCTTATGCTATGCTACCA	1361
Qy	433 GTGCTGCTCCCTATGGGATGTGCGCAGCTTCTTCAAGGCTCTGGGCACTGGAGGCTGA	492
Db	1362 GTGCTGCTCCCTATGGGATGTGCGCAGCTTCTTCAAGGCTCTGGGCACTGGAGGCTGA	1421
Qy	493 AGACCTTCACTTGATGATGAGGAGTCTTCAAAAAGGCGCTTGGGCTCTCTTGGCAGACA	552
Db	1422 AGACCTTCACTTGATGATGAGGAGTCTTCAAAAAGGCGCTTGGGCTCTCTTGGCAGACA	1481
Qy	553 AGCAGAGAACCACTATGACAGGACCTGGATGAGCTCCAGCTGGAGATGAGGAGCTCAAA	612
Db	1482 AGCAGAGAACCACTATGACAGGACCTGGATGAGCTCCAGCTGGAGATGAGGAGCTCAAA	1541
Qy	613 GCCACACCCAGTGTCCAGTGTAGCCCTACTTCCAGGCGCCCTTCAAGCCCTGTGAGTACT	672
Db	1542 GCCACACCCAGTGTCCAGTGTAGCCCTACTTCCAGGCGCCCTTCAAGCCCTGTGAGTACT	1601
Qy	673 CTTTGAAGCTGGGGCATCCGCTGGCGGTGTGGGCCATCGTGTGCTCTCCGTGCTCTG	732
Db	1602 CTTTGAAGCTGGGGCATCCGCTGGCGGTGTGGGCCATCGTGTGCTCTCCGTGCTCTG	1661
Qy	733 CAATGGACTGTGCTGTGCTGACCGTGTTCGCTGGGGGCTGCGCCCTGCGCCCGGTCAA	792
Db	1662 CAATGGACTGTGCTGTGCTGACCGTGTTCGCTGGGGGCTGCGCCCTGCGCCCGGTCAA	1721

Qy	793 GTTTGTGTAGTGGATTCGAGGCGCAACACCTTGACTGACATTTCTGTGGCTTCT 852	
Db	1722 GTTTGTGTAGTGGATTCGAGGCGCAACACCTTGACTGACATTTCTGTGGCTTCT 1781	
Qy	853 AGCCTCAGTCAGTGGCTTGGTTCAGTTCCTGAGTACGAGGCGGCTGGAGAC 912	
Db	1782 AGCCTCAGTCAGTGGCTTGGTTCAGTTCCTGAGTACGAGGCGGCTGGAGAC 1841	
Qy	913 GGGGCTAGGCTGCGGGGCTTCTTGGCAGTACTTGGTGGAGGATCGGTGCT 972	
Db	1842 GGGGCTAGGCTGCGGGGCTTCTTGGCAGTACTTGGTGGAGGATCGGTGCT 1901	
Qy	973 GCTGCTACTCTGGCCGAGTGCAGTGCAGGCTCTCCGTCTCTGTGTCCGGGCTATGG 1032	
Db	1902 GCTGCTACTCTGGCCGAGTGCAGTGCAGGCTCTCCGTCTCTGTGTCCGGGCTATGG 1961	
Qy	1033 GAAATCCCTCCCTGGGACAGCTTCGAGCAGGGGCTCTAGGCTGCTGGCAGTGGCAGG 1092	
Db	1962 GAAATCCCTCCCTGGGACAGCTTCGAGCAGGGGCTCTAGGCTGCTGGCAGTGGCAGG 2021	
Qy	1093 GCTGGCCGCGCACTGCCCCCTGGCTCAGTGGGAGATACGGGGCTTCCCACTCTGCT 1152	
Db	2022 GCTGGCCGCGCACTGCCCCCTGGCTCAGTGGGAGATACGGGGCTTCCCACTCTGCT 2081	
Qy	1153 GCTTACGCGCCACTGAGGGTTCAGCCAGCAGCCCTTGGGCTTTCACCGTGGGCTTGGTAT 1212	
Db	2082 GCTTACGCGCCACTGAGGGTTCAGCCAGCAGCCCTTGGGCTTTCACCGTGGGCTTGGTAT 2141	
Qy	1213 GATGAATCTCTTCTGTTTCTGCTGGTGGGCGGCTTACATCAAACTGTACTGTGACT 1272	
Db	2142 GATGAATCTCTTCTGTTTCTGCTGGTGGGCGGCTTACATCAAACTGTACTGTGACT 2201	
Qy	1273 GCGCGGGGCACTTTGAGGCGGCTGGGACTGGCCATGTGAGGACGTGGCTGCT 1332	
Db	2202 GCGCGGGGCACTTTGAGGCGGCTGGGACTGGCCATGTGAGGACGTGGCTGCT 2261	
Qy	1333 CATCTTCGACAGCGGCTCTCTACTGTCCGTGGGCTTCTCTAGCTTCGCTTCCATGCT 1392	
Db	2262 CATCTTCGACAGCGGCTCTCTACTGTCCGTGGGCTTCTCTAGCTTCGCTTCCATGCT 2321	
Qy	1393 GGGCTCTTCTCTGTCAGGCGGCGGTCAAGTCTGTCTCTGTGTGTGTGTGTGTGTGTGT 1452	
Db	2322 GGGCTCTTCTCTGTCAGGCGGCGGTCAAGTCTGTCTCTGTGTGTGTGTGTGTGTGTGT 2381	
Qy	1453 GCCTGCTGCTCAACCCACTGCTACTCTTCAACCCCTTCCGCGGATGACT 1512	
Db	2382 GCCTGCTGCTCAACCCACTGCTACTCTTCAACCCCTTCCGCGGATGACT 2441	
Qy	1513 TCGCGGGCTTCGGGCGCGGAGGAGTCAAGGCGGCTTACAGGCTTATGCTGGGCGGGGA 1572	
Db	2442 TCGCGGGCTTCGGGCGCGGAGGAGTCAAGGCGGCTTACAGGCTTATGCTGGGCGGGGA 2501	
Qy	1573 GCTGGAGAGAGCTCTGTGATTTACCCAGGCGCTGTGTGTGTGTGTGTGTGTGTGTGTGT 1632	
Db	2502 GCTGGAGAGAGCTCTGTGATTTACCCAGGCGCTGTGTGTGTGTGTGTGTGTGTGTGT 2561	
Qy	1633 CATTCGGAAGCTTCTGAAGCTGGGCGGCTTGGGCTGGAGACCTATGCTTCCCTC 1692	
Db	2562 CATTCGGAAGCTTCTGAAGCTGGGCGGCTTGGGCTGGAGACCTATGCTTCCCTC 2621	
Qy	1693 AGTGACCTCATCTCTCTGTCAGCAGCCAGGGGCGGCTTGGAGGCGGAGCCATTTGTGT 1752	
Db	2622 AGTGACCTCATCTCTCTGTCAGCAGCCAGGGGCGGCTTGGAGGCGGAGCCATTTGTGT 2681	
Qy	1753 AGAGCCAGAGGGGAAACCACTTTTGGGAAACCCCAACCTCTCATAGGATGAGAACTGTGCT 1812	
Db	2682 AGAGCCAGAGGGGAAACCACTTTTGGGAAACCCCAACCTCTCATAGGATGAGAACTGTGCT 2741	
Qy	1813 GAGGCGAGGGATCTAGCCAGGCTGGAGGCTTGTTCAGGGGCTGGGCTTTCAGCC 1872	
Db	2742 GAGGCGAGGGATCTAGCCAGGCTGGAGGCTTGTTCAGGGGCTGGGCTTTCAGCC 2801	

QY 1873 CTCTGGCTTGGCTTGTGCTTACACGTTAAATATCCCTCCCACTTCTTCTTCCCTC 1932
DB |||||
2802 CTCTGGCTTGGCTTGTGCTTACACGTTAAATATCCCTCCCACTTCTTCTTCCCTC 2861
QY 1933 TCTTCCCTTCTCTCTCCCTCGGTGAATGATGGCTGCTTCTTAAACAATAACAACA 1992
DB |||||
2862 TCTTCCCTTCTCTCTCCCTCGGTGAATGATGGCTGCTTCTTAAACAATAACAACA 2921
QY 1993 AAATCAGCAGTGTGATCTATAGCAGGATGGCCAGTACCTGGCTCCACTGATCACCTCT 2052
DB |||||
2922 AAATCAGCAGTGTGATCTATAGCAGGATGGCCAGTACCTGGCTCCACTGATCACCTCT 2981
QY 2053 CTCTGTGACCATCACCAAGGGTCCCTTCTGGCTGGCTTCTTCCCTTGGCTTCTCTCAGC 2112
DB |||||
2982 CTCTGTGACCATCACCAAGGGTCCCTTCTGGCTGGCTTCTTCCCTTGGCTTCTCTCAGC 3041
QY 2113 TTCACTTGTATCTGGGCTCTTCTTCTGTCATGCTGAAGCTGTGGACCAAGACCTTGA 2172
DB |||||
3042 TTCACTTGTATCTGGGCTCTTCTTCTGTCATGCTGAAGCTGTGGACCAAGACCTTGA 3101
QY 2173 CTTTGTCTGCTTAAGGAAATGAGGAAGTAAGACAGTGAAGGGTGGAGGTTGATC 2232
DB |||||
3102 CTTTGTCTGCTTAAGGAAATGAGGAAGTAAGACAGTGAAGGGTGGAGGTTGATC 3161
QY 2233 AGGGCAGTGGACAGGGAGACCTCACAGAGAAAGGCTGGAAGTGATTTCCCGTGTGA 2292
DB |||||
3162 AGGGCAGTGGACAGGGAGACCTCACAGAGAAAGGCTGGAAGTGATTTCCCGTGTGA 3221
QY 2293 CTATGTAGTAGGATACAAATGTGTTCATGTACCATTAATCTTTGACATATGCCATGCAT 2352
DB |||||
3222 CTATGTAGTAGGATACAAATGTGTTCATGTACCATTAATCTTTGACATATGCCATGCAT 3281
QY 2353 AAAGACTTCTTAAATAAGCTTTGGAAGAGATTAATAAAAA 2396
DB |||||
3282 AAAGACTTCTTAAATAAGCTTTGGAAGAGATTAATAAAAA 3325

RESULT 14

AD121981
ID AD121981 standard; cDNA; 2964 BP.
XX
AC AD121981;
XX
XX 15-APR-2004 (first entry)
XX
XX Novel human protein cDNA #240.
DE
DE forensic; nutritional source; damaged tissue; diseased tissue;
KW myeloid cell disorder; lymphoid cell disorder;
KW bone cartilage tissue growth; tendon tissue growth;
KW ligament tissue growth; nerve tissue growth; regeneration; wound healing;
KW tissue repair; tissue replacement; burn; incision; ulcer; cancer; human;
KW ss; gene.
XX
OS Homo sapiens.
XX
XX WO2003025148-A2.
XX
XX 27-MAR-2003.
XX
XX 19-SEP-2002; 2002WO-US029964.
XX
XX 19-SEP-2001; 2001US-0323739P.
XX
XX 13-SEP-2002; 2002US-00323739.
XX
XX (HYSE-) HYSEQ INC.
XX
XX Tang YT, Asundi V, Goodrich RW, Ren F, Zhang J, Zhao QA, Wang J;
PI Ghosh M, Xue AJ, Wehrman T, Weng G, Zhou P, Drmanac R, Wang D;
PI Haley-Vicente D;
XX
XX WPI; 2003-354603/33.
DR P-PSDB; AD121265.

XX New polynucleotides and secreted proteins, useful for treating myeloid or
PT lymphoid cell disorders, in bone cartilage, tendon, ligament and nerve
PT tissue growth or regeneration, in wound healing, and in tissue repair and
PT replacement.
XX
PS Claim 1; SEQ ID NO 240; 156pp; English.
XX
CC The invention relates to an isolated polynucleotide encoding a
CC polypeptide with biological activity. The polynucleotides and
CC polypeptides are useful in diagnostics, forensics, gene mapping,
CC identification of mutations responsible for genetic disorders and other
CC traits, to assess biodiversity, as nutritional sources or supplements.
CC The polynucleotides may also be used as molecular weight markers.
CC chromosome markers or map related gene positions, or as an antigen to
CC raise anti-DNA antibodies or elicit immune response. The polypeptides are
CC useful for raising antibodies, as markers for tissues in which the
CC corresponding polypeptide is expressed, for re-engineering damaged or
CC diseased tissues, for treating myeloid or lymphoid cell disorders, in
CC bone cartilage, tendon, ligament and/or nerve tissue growth or
CC regeneration, in wound healing, in tissue repair and replacement, in
CC healing of burns, incisions and ulcers, and in treating cancer. The
CC present sequence represents cDNA encoding a novel human protein.
XX
SQ Sequence 2964 BP; 563 A; 936 C; 825 G; 640 T; 0 U; 0 Other;
Query Match 90.2%; Score 2243; DB 10; Length 2964;
Best Local Similarity 99.7%; Pred. No. 0;
Matches 2257; Conservative 0; Mismatches 5; Indels 1; Gaps 1;
QY 133 AGCGGCTTCCAAACACACACCGCATCTGGGAAATTTGGAGCTTGACACCTTCAGCCAGCTGAG 192
DB 703 ATCGGCTTCCAAACACACCGCATCTGGGAAATTTGGAGCT-GACACCTTCAGCCAGCTGAG 761
QY 193 CTCCTGCAAGCCCTGGATCTTAGCTGGAAACGCGCATCCGGTCCATCCACCTGAGGCTT 252
DB 762 CTCCTGCAAGCCCTGGATCTTAGCTGGAAACGCGCATCCGGTCCATCCACCTGAGGCTT 821
QY 253 CTCACCCCTGCATCCCTGGTCAAGCTGGAGCTTGACAGACCAACCGCTGACACCTGACACCTGCC 312
DB 822 CTCACCCCTGCATCCCTGGTCAAGCTGGAGCTTGACAGACCAACCGCTGACACCTGCC 881
QY 313 CTGGCTGGACTTGGGGCTTGTATGATCTGAAGCTCAAAAGGGAACCTTGCTCTCTCCCA 372
DB 882 CTGGCTGGACTTGGGGCTTGTATGATCTGAAGCTCAAAAGGGAACCTTGCTCTCTCCCA 941
QY 373 GGCTTCTCCAAGGACAGTTTCCCAAACTGAGAGATCTCGAGGTGCTTATGCTTACCA 432
DB 942 GGCTTCTCCAAGGACAGTTTCCCAAACTGAGAGATCTCGAGGTGCTTATGCTTACCA 1001
QY 433 GTGCTGTCTTATGGATGTGTCAGCTTCTTCAAGGCTCTTGGGAGTGGGAGGCTGA 492
DB 1002 GTGCTGTCTTATGGATGTGTCAGCTTCTTCAAGGCTCTTGGGAGTGGGAGGCTGA 1061
QY 493 AGACCTTTCACCTTGATGATGAGGATCTTCAAAAGGCTCTTGGGCTCTTGGCAGACA 552
DB 1062 AGACCTTTCACCTTGATGAGGATCTTCAAAAGGCTCTTGGGCTCTTGGCAGACA 1121
QY 553 AGCAGAGAACCACTATGACCAAGGACCTGGATGAGCTCCAGCTGGAGATGGAGGACTCAAA 612
DB 1122 AGCAGAGAACCACTATGACCAAGGACCTGGATGAGCTCCAGCTGGAGATGGAGGACTCAAA 1181
QY 613 GCCACACCCAGTGTCCAGTGTAGCCCTACTCCAGGCCCCCTTCAGGCCCTTGTGAGTACCT 672
DB 1182 GCCACACCCAGTGTCCAGTGTAGCCCTACTCCAGGCCCCCTTCAGGCCCTTGTGAGTACCT 1241
QY 673 CTTTGAAGCTGGGGCATCCGCTGGCCGTGTGGGCCATCGTGTGTCTCCGTCCTCTG 732
DB 1242 CTTTGAAGCTGGGGCATCCGCTGGCCGTGTGGGCCATCGTGTGTCTCCGTCCTCTG 1301
QY 733 CAATGGAATGCTGCTGACCGTGTTCGCTGGGGGCTGCCCCCTGCCCCCGGTCAA 792
DB 1302 CAATGGAATGCTGCTGACCGTGTTCGCTGGGGGCTGCCCCCTGCCCCCGGTCAA 1361

Qy	1873	CTCTGGGTTGGCCCTTCTTTCACACAGTGTAATAATCCCTCGCCCATTTCTTCTCTTCGCCCTC	1932
Db	2442	CTCTGGGTTGGCCCTTCTTTCACACAGTGTAATAATCCCTCGCCCATTTCTTCTCTTCGCCCTC	2501
Qy	1933	TCCTTCCCTTTCCTCTCTCCCTCCGCTGGAATGATGCTGCTTCTTAAACAAATACAAACCA	1992
Db	2502	TCCTTCCCTTTCCTCTCTCCCTCCGCTGGAATGATGCTGCTTCTTAAACAAATACAAACCA	2561
Qy	1993	AAACTCAGCAGTGTGATCTATAGCAGGATGGCCACAGTACCTGGCTCCACTGATCACCTCT	2052
Db	2562	AAACTCAGCAGTGTGATCTATAGCAGGATGGCCACAGTACCTGGCTCCACTGATCACCTCT	2621
Qy	2053	CTCCTGTGACCATCACAAACGGGTGGCTCTTGGCCCTGGCTTTCCCTTGGCCCTTCTCAGC	2112
Db	2622	CTCCTGTGACCATCACAAACGGGTGGCTCTTGGCCCTGGCTTTCCCTTGGCCCTTCTCAGC	2681
Qy	2113	TTCCACCTTGATACCTGGGCGCTCTTCCCTTGTCTGTCATGCTCTGAAAGCTGTGGACCAAGACACCTGGA	2172
Db	2682	TTCCACCTTGATACCTGGGCGCTCTTCCCTTGTCTGTCATGCTCTGAAAGCTGTGGACCAAGACACCTGGA	2741
Qy	2173	CTTTTGTCTGCTTAAAGGGAATGAGGGAAGTAAAGACAGTGAAGGGGTGGAGGGTTGATC	2232
Db	2742	CTTTTGTCTGCTTAAAGGGAATGAGGGAAGTAAAGACAGTGAAGGGGTGGAGGGTTGATC	2801
Qy	2233	AGGGCACAGTGGACAGGGAGACCTCAACAGAAAGCCCTGGAAAGTGAATTTCCCGTGTGA	2292
Db	2802	AGGGCACAGTGGACAGGGAGACCTCAACAGAAAGCCCTGGAAAGTGAATTTCCCGTGTGA	2861
Qy	2293	CTCATGGATAGGATACAAAATGTGTTCCATGTACCATTTAATCTTGACATATGCCATGCAT	2352
Db	2862	CTCATGGATAGGATACAAAATGTGTTCCATGTACCATTTAATCTTGACATATGCCATGCAT	2921
Qy	2353	AAAGACTTCTCTATTAAATAAGCTTTTGGAAAGAGATTAAAAAAA	2395
Db	2922	AAAGACTTCTCTATTAAATAAGCTTTTGGAAAGAGATTAAAAAAA	2964
RESULT 15			
AAD37674			
ID	AAD37674 standard; cDNA; 3429 BP.		
XX	AAD37674;		
XX	27-AUG-2002 (first entry)		
XX	Human G-protein coupled receptor-10 (GCREC-10) cDNA.		
XX	Human; G-protein coupled receptor-10; GCREC-10; atherosclerosis; cancer;		
KW	cell proliferative disorder; gastrointestinal; autoimmune; metabolic;		
KW	neurological; inflammatory; cardiovascular; viral infection; anorexia;		
KW	cirrhosis; multiple sclerosis; Huntington's disease; gene therapy;		
KW	Alzheimer's disease; heart failure; hepatitis; cholecystitis; obesity;		
KW	rheumatoid arthritis; Hashimoto's thyroiditis; angina pectoris;		
KW	osteoporosis; gene; ss.		
XX	Homo sapiens.		
XX	Key Location/Qualifiers		
FT	CDS	1..2847	
FT		/*tag= a	
FT		/product= "Human GCREC-10 protein"	
FT	sig_peptide	1..75	
FT		/*tag= b	
FT	mat_peptide	76..2844	
FT		/*tag= c	
FT		/product= "Mature GCREC-10 protein"	
XX	WO200226825-A2.		
PN			
XX	04-APR-2002.		
PD			
XX	28-SEP-2001; 2001WO-US030661.		
XX			
XX			


```
Db 2577 CATTCGGAAGCTTCTGAAGCTGGGGCGGCCCTCGGCTGGAGACCTATGGCTTCCCTC 2636
QY 1693 AGTGACCTCTATCTCCTGTGACAGCCAGGGGCCCCCAGGCTGGAGGCGAGCCATTGTGT 1752
Db 2637 AGTGACCTCTATCTCCTGTGACAGCCAGGGGCCCCCAGGCTGGAGGCGAGCCATTGTGT 2696
QY 1753 AGAGCCAGAGGGGAGCCACTTTGGGAAACCCCAACCCCTCCATGGATGGAGAACTGCTGCT 1812
Db 2697 AGAGCCAGAGGGGAGCCACTTTGGGAAACCCCAACCCCTCCATGGATGGAGAACTGCTGCT 2756
QY 1813 GAGGGCAGAGGGATCTACCCAGCAGGTGGAGGCTGTTCAGGGGTGGGGCTTTCAGCC 1872
Db 2757 GAGGGCAGAGGGATCTACCCAGCAGGTGGAGGCTGTTCAGGGGTGGGGCTTTCAGCC 2816
QY 1873 CTCTGGCTTGGCTTTGCTTTCACAGTGTAAATATCCCTCCCAATTCCTCTCTCCCTC 1932
Db 2817 CTCTGGCTTGGCTTTGCTTTCACAGTGTAAATATCCCTCCCAATTCCTCTCTCCCTC 2876
QY 1933 TCTTCCCTTCTCTCTCCCTCGGTGAATGATGGCTGCTTCTAAACAAATACAACA 1992
Db 2877 TCTTCCCTTCTCTCTCCCTCGGTGAATGATGGCTGCTTCTAAACAAATACAACA 2936
QY 1993 AAACCTCAGCAGTGATCTATAGCAGGATGCCAGTACCTGGCTCCACTGATCACTCT 2052
Db 2937 AAACCTCAGCAGTGATCTATAGCAGGATGCCAGTACCTGGCTCCACTGATCACTCT 2996
QY 2053 CTCCTGTGACCATCACCAACGGGTGCTCTGGCTGGCTTTCCTTGGCTTCTCTCAGC 2112
Db 2997 CTCCTGTGACCATCACCAACGGGTGCTCTGGCTGGCTTTCCTTGGCTTCTCTCAGC 3056
QY 2113 TTCACCTTGATACTGGGCTCTTCCTTGTCTGTCTGAAGCTGTGGACGAGACCTGGA 2172
Db 3057 TTCACCTTGATACTGGGCTCTTCCTTGTCTGTCTGAAGCTGTGGACGAGACCTGGA 3116
QY 2173 CTTTGTCTGCTTAAGGGAATGAGGGAAGTAAAGACAGTGAAGGGTGGAGGTTGATC 2232
Db 3117 CTTTGTCTGCTTAAGGGAATGAGGGAAGTAAAGACAGTGAAGGGTGGAGGTTGATC 3176
QY 2233 AGGGCACAGTGGACAGGGAGACCTCACAGAGAAAGCCTGGAGGTGATTTCCCGTGTGA 2292
Db 3177 AGGGCACAGTGGACAGGGAGACCTCACAGAGAAAGCCTGGAGGTGATTTCCCGTGTGA 3236
QY 2293 CTCATGGATAGGATACAAATGTGTTCATGTACCATTAATCTTGACATATGCCATGCAT 2352
Db 3237 CTCATGGATAGGATACAAATGTGTTCATGTACCATTAATCTTGACATATGCCATGCAT 3296
QY 2353 AAAGACTTCCCTATTAAATTAAGCTTTGGNAGAGATTAAAA 2393
Db 3297 AAAGACTTCCCTATTAAATTAAGCTTTGGNAGAGATTACACA 3337
```

Search completed: April 14, 2005, 12:29:57
Job time : 885.045 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 14, 2005, 12:03:58 ; Search time 943.376 Seconds
(without alignments)
15987.096 Million cell updates/sec

Title: US-10-664-667-4

Perfect score: 2486

Sequence: 1 taatcagactcactatagg.....agtgacacctaattcaatt 2486

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 5622541 seqs, 303355566 residues

Total number of hits satisfying chosen parameters: 11245082

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications NA.*
1: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq.*
2: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq.*
3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq.*
4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq.*
5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq.*
6: /cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq.*
7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq.*
8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq.*
9: /cgn2_6/ptodata/1/pubpna/US09A_PUBCOMB.seq.*
10: /cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq.*
11: /cgn2_6/ptodata/1/pubpna/US09C_PUBCOMB.seq.*
12: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq.*
13: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq.*
14: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq.*
15: /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq.*
16: /cgn2_6/ptodata/1/pubpna/US10D_PUBCOMB.seq.*
17: /cgn2_6/ptodata/1/pubpna/US10E_PUBCOMB.seq.*
18: /cgn2_6/ptodata/1/pubpna/US10F_PUBCOMB.seq.*
19: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq.*
20: /cgn2_6/ptodata/1/pubpna/US11_NEW_PUB.seq.*
21: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq.*
22: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2485	100.0	2486	10	US-09-851-595-4
2	2485	100.0	2486	17	US-10-664-667-4
3	2257.4	90.8	2711	10	US-09-851-595-7
4	2257.4	90.8	2711	17	US-10-664-667-7
5	2255.8	90.7	2988	17	US-10-302-172-723
6	2247.8	90.4	3349	14	US-10-270-336-4
7	2247.2	90.4	3325	17	US-10-331-496A-52
8	2241	90.1	3429	18	US-10-398-036-26
9	2238	90.0	3443	17	US-10-331-496A-94
10	2238	90.0	3492	10	US-09-851-595-10
11	2238	90.0	3492	17	US-10-664-667-10
					Sequence 4, Appli
					Sequence 4, Appli
					Sequence 7, Appli
					Sequence 723, App
					Sequence 4, Appli
					Sequence 52, Appl
					Sequence 26, Appl
					Sequence 94, Appl
					Sequence 10, Appl

12	2238	90.0	3492	18	US-10-737-450-31	Sequence 31, Appl
13	2231.4	89.8	3273	17	US-10-295-027-1119	Sequence 1119, Ap
14	2231.4	89.8	3273	18	US-10-783-528-43	Sequence 43, Appl
15	2219.6	89.3	3381	10	US-09-970-944-5	Sequence 5, Appli
16	2085.8	83.9	3119	15	US-10-225-567A-580	Sequence 580, App
17	1898	76.3	1899	10	US-09-851-595-6	Sequence 6, Appli
18	1898	76.3	1899	17	US-10-664-667-6	Sequence 6, Appli
19	1754.4	70.6	2208	10	US-09-851-595-9	Sequence 9, Appli
20	1754.4	70.6	2208	17	US-10-664-667-9	Sequence 9, Appli
21	1754.4	70.6	2901	10	US-09-851-595-12	Sequence 12, Appl
22	1754.4	70.6	2901	17	US-10-664-667-12	Sequence 12, Appl
23	1754.4	70.6	3438	14	US-10-176-847-89	Sequence 89, Appl
24	1752.6	70.5	2487	14	US-10-270-336-1	Sequence 1, Appli
25	1743	70.1	119596	14	US-10-270-336-3	Sequence 3, Appli
26	1383.4	55.6	3637	10	US-09-851-595-1	Sequence 1, Appli
27	1383.4	55.6	3637	17	US-10-664-667-1	Sequence 1, Appli
28	1336	53.7	2901	10	US-09-851-595-3	Sequence 3, Appli
29	1336	53.7	2901	17	US-10-664-667-3	Sequence 3, Appli
30	983.4	39.6	1078	17	US-10-264-049-395	Sequence 395, App
31	723.2	29.1	1173	17	US-10-343-650A-65	Sequence 65, Appl
32	544.8	21.9	622	17	US-10-172-118-2496	Sequence 2496, Ap
33	544.8	21.9	622	17	US-10-342-887-2496	Sequence 2496, Ap
34	439.4	17.7	453	9	US-09-876-889-240	Sequence 240, App
35	416.2	16.7	422	10	US-09-918-995-34626	Sequence 34626, A
36	356.2	14.3	2082	19	US-10-851-470-3	Sequence 3, Appli
37	355.6	14.3	2651	17	US-10-295-027-483	Sequence 483, App
38	355.6	14.3	2651	17	US-10-295-027-1113	Sequence 1113, Ap
39	355.6	14.3	2651	17	US-10-173-999-27	Sequence 27, Appl
40	355.6	14.3	2724	15	US-10-251-385-263	Sequence 263, App
41	355.6	14.3	2724	15	US-10-225-567A-421	Sequence 421, App
42	355.6	14.3	2724	17	US-10-174-456-4	Sequence 4, Appli
43	355.6	14.3	2724	18	US-10-751-736-21	Sequence 21, Appl
44	355.6	14.3	2880	17	US-10-295-027-1114	Sequence 1114, Ap
45	355.6	14.3	2880	19	US-10-482-029-157	Sequence 157, App

ALIGNMENTS

RESULT 1

US-09-851-595-4
; Sequence 4, Application US/09851595
; Publication No. US20030166047A1
; GENERAL INFORMATION:
; APPLICANT: Gu, Wei
; TITLE OF INVENTION: NOVEL G-PROTEIN COUPLED RECEPTORS AND USES THEREFOR
; FILE REFERENCE: MNI-080CP
; CURRENT APPLICATION NUMBER: US/09/851,595
; CURRENT FILING DATE: 2000-05-08
; PRIOR APPLICATION NUMBER: 09/556,588
; PRIOR FILING DATE: 2000-05-08
; PRIOR APPLICATION NUMBER: 60/132,896
; PRIOR FILING DATE: 1999-05-06
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 4
; LENGTH: 2486
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (2)..(1900)
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (172)
; OTHER INFORMATION: n = any nucleotide
US-09-851-595-4

Query Match 100.0%; Score 2485; DB 10; Length 2486;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2486; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 TAATACGACTCACTATAGGAAAGCTGGTACGCTCAGGTAACCGGTCGGAATTCGCG 60

Db	1	TAATACGACTC	ATATGGGAAAGCTGGT	AGCCCTGCGAGTACCGG	TCCGGAATTC	CCCGG	60
Qy	61	GTGACCCACAGCGT	CCGTTGGAGCGGAGCAGGGT	CTGAGCCTGCGGGCTCAT	CCAGCGCTC	120	
Db	61	GTGACCCACAGCGT	CCGTTGGAGCGGAGCAGGGT	CTGAGCCTGCGGGCTCAT	CCAGCGCTC	120	
Qy	121	TC TTGCTGCCCT	TAGCGGCTCCAA	CAACCGCATCTGGGAAATTGGAGCTNGACACCTT	180		
Db	121	TC TTGCTGCCCT	TAGCGGCTCCAA	CAACCGCATCTGGGAAATTGGAGCTNGACACCTT	180		
Qy	181	CAGCCAGCTGAGCT	CCCTGGAAGCCCTGGAT	TCTAGCTGGAAACGCGCATTCGGTTCATCCA	240		
Db	181	CAGCCAGCTGAGCT	CCCTGGAAGCCCTGGAT	TCTAGCTGGAAACGCGCATTCGGTTCATCCA	240		
Qy	241	CCCTGAGCGCTT	CTCCACCCCTGCACT	CCCTGGTCAAGCTGACCTGCAGACAACCAAGCT	300		
Db	241	CCCTGAGCGCTT	CTCCACCCCTGCACT	CCCTGGTCAAGCTGACCTGCAGACAACCAAGCT	300		
Qy	301	GACCACACTGCCC	TGGCTGGACTTTGGGGCTTGATGCAT	CTGAAGCTCAAAAGGGAACCT	360		
Db	301	GACCACACTGCCC	TGGCTGGACTTTGGGGCTTGATGCAT	CTGAAGCTCAAAAGGGAACCT	360		
Qy	361	TGCTCTCTCC	AGGCTTCTCCAAAGACAGTTT	TCCAAAACTGAGGATCTTGGAGGTGCC	420		
Db	361	TGCTCTCTCC	AGGCTTCTCCAAAGACAGTTT	TCCAAAACTGAGGATCTTGGAGGTGCC	420		
Qy	421	TTATGCGCTAC	CAGTGCTGGCTAT	GGGATGTGTCAGCTTCTCAAGGCTCTGGGCA	480		
Db	421	TTATGCGCTAC	CAGTGCTGGCTAT	GGGATGTGTCAGCTTCTCAAGGCTCTGGGCA	480		
Qy	481	GTGGAGGCTGA	AGACTTCACTTTGATGATGAGGAGTCTT	CAAAAAGGCCCTTGGGCT	540		
Db	481	GTGGAGGCTGA	AGACTTCACTTTGATGATGAGGAGTCTT	CAAAAAGGCCCTTGGGCT	540		
Qy	541	CCTTGCCAGACA	AGAGAACCACTATGACAGGACCTG	ATGAGCTCCAGCTGGAGAT	600		
Db	541	CCTTGCCAGACA	AGAGAACCACTATGACAGGACCTG	ATGAGCTCCAGCTGGAGAT	600		
Qy	601	GGAGGACTCAA	AGCCACACCCAGTGTCCAGTGTAGCCCTACT	CCAGGCCCTTCAAGCC	660		
Db	601	GGAGGACTCAA	AGCCACACCCAGTGTCCAGTGTAGCCCTACT	CCAGGCCCTTCAAGCC	660		
Qy	661	CTGTGAGTACCT	CTTTGAAAGCTGGGGCATCCGCT	GTGGCCGTGTGGGCACTCGTGTGCT	720		
Db	661	CTGTGAGTACCT	CTTTGAAAGCTGGGGCATCCGCT	GTGGCCGTGTGGGCACTCGTGTGCT	720		
Qy	721	CTCCGTGCTCT	GCAATGAGACTGGTGTGCTGACCGTGTTCGCT	GTGGCGGCTTGC	CCCGCT	780	
Db	721	CTCCGTGCTCT	GCAATGAGACTGGTGTGCTGACCGTGTTCGCT	GTGGCGGCTTGC	CCCGCT	780	
Qy	781	GCCCCCGGTCA	AGTTTGTGTAGGTGCGATTG	CGAGGCGCAACACTTGA	CTGCGCATTTTC	840	
Db	781	GCCCCCGGTCA	AGTTTGTGTAGGTGCGATTG	CGAGGCGCAACACTTGA	CTGCGCATTTTC	840	
Qy	841	CTGTGGCTTCT	TAGCCTCAGTCGATGCCCTGACCTTTTGGT	CTCAGTCTCTGAGTAGCGGAGC	900		
Db	841	CTGTGGCTTCT	TAGCCTCAGTCGATGCCCTGACCTTTTGGT	CTCAGTCTCTGAGTAGCGGAGC	900		
Qy	901	CCGCTGGGAGAC	GGGGCTTAGGCTGCGCGGCCACTCGGCTT	CCTGGCAGTACTTGGGTCCGA	960		
Db	901	CCGCTGGGAGAC	GGGGCTTAGGCTGCGCGGCCACTCGGCTT	CCTGGCAGTACTTGGGTCCGA	960		
Qy	961	GGCATTCGGTGT	GTGTCTACTCTGGCGCAGTGCAGTGCAGCT	TCCCGTCTCTGTGT	1020		
Db	961	GGCATTCGGTGT	GTGTCTACTCTGGCGCAGTGCAGTGCAGCT	TCCCGTCTCTGTGT	1020		
Qy	1021	CCGGGCTATGGG	AAAGTCCCGCTCCCTGGGAGAGTTCGAGCAGGGGT	CTTAGGCTGCCT	1080		
Db	1021	CCGGGCTAT	GGGAAAGTCCCGCTCCCTGGGAGAGTTCGAGCAGGGGT	CTTAGGCTGCCT	1080		
Qy	1081	GGCACTGGCAGGG	GTGGCGCGCACTGCCCCCTGGGCTT	CTAGTGGGAGAAATACGGGGGCTC	1140		

Db	1081	GGCACTGGCAGGGGTGGCCGCGGACATGCCCCCTGGCCCTCAGTGGGAGAAATACGGGGCGCTC	1141
Qy	1141	CCCACTCTGCTGCCCTCACGCGCACCTGAGGGTCAGCCAGACAGCCCTGGGCTTCACCGT	1200
Db	1141	CCCACTCTGCTGCCCTACGCGGCACTGAGGGTCAGCCAGACAGCCCTGGGCTTCACCGT	1200
Qy	1201	GGCCCTGGTGATGATGAATCTCTTCTGTTCCTGGTCGTGGCCGGTCCCTACATCAAACT	1260
Db	1201	GGCCCTGGTGATGATGAATCTCTTCTGTTCCTGGTCGTGGCCGGTCCCTACATCAAACT	1260
Qy	1261	GTACTGTGACTGCGCGGGGGGACATTTGAGGCGCGTGGGACTGGCCCATGGTAGGGCA	1320
Db	1261	GTACTGTGACTGCGCGGGGGGACATTTGAGGCGCGTGGGACTGGCCCATGGTAGGGCA	1320
Qy	1321	CGTGGCCTGGCTCATCTTTCCAGACGGGCTCTACTGTGCCGTGGCCCTTCCTCAGCTT	1380
Db	1321	CGTGGCCTGGCTCATCTTCCAGACGGGCTCTCTACTGTGCCGTGGCCCTTCCTCAGCTT	1380
Qy	1381	CGCCTCCATGCTGGGCGCTTCTCCCTGTGCAGCGCCGAGGCGGTCAAGTCTGTCTGTGTGT	1440
Db	1381	CGCCTCCATGCTGGGCGCTTCTCCCTGTGCAGCGCCGAGGCGGTCAAGTCTGTCTGTGTGT	1440
Qy	1441	GGTGTCCCCCTGGCTGCTGCTCAACCCATCTGCTGATCTGTGCTCTTCAAACCCCACTT	1500
Db	1441	GGTGTCCCCCTGGCTGCTGCTCAACCCATCTGCTGATCTGTGCTCTTCAAACCCCACTT	1500
Qy	1501	CCGGGATGACTTTCCGGCGGCTTCGGCCCGCGCAGGGGACTCAGGGGCCCTAGGCTATGC	1560
Db	1501	CCGGGATGACTTTCCGGCGGCTTCGGCCCGCGCAGGGGACTCAGGGGCCCTAGGCTATGC	1560
Qy	1561	TGCGGCCGGGGAGCTGGAGAAGAGCTCTGTGATTTCAACCAGGCCCTGGTAGGCTTCTC	1620
Db	1561	TGCGGCCGGGGAGCTGGAGAAGAGCTCTGTGATTTCAACCAGGCCCTGGTAGGCTTCTC	1620
Qy	1621	TGATGTGGATCTCATTTCTGGAAGCTTCTGAAGCTGGCGGCGCCCTGGGCTGGAGACCTA	1680
Db	1621	TGATGTGGATCTCATTTCTGGAAGCTTCTGAAGCTGGCGGCGCCCTGGGCTGGAGACCTA	1680
Qy	1681	TGGCTTCCCCTCAGTGACCTCATCTCTGTCTCAGCAGCCAGGGGCGCCAGGGCTGGAGG	1740
Db	1681	TGGCTTCCCCTCAGTGACCTCATCTCTGTCTCAGCAGCCAGGGGCGCCAGGGCTGGAGG	1740
Qy	1741	CAGCCATTGTGAGAGCCAGAGGGGGAACCACTTTTGGGAACCCCCAACCCCTCCATGGATGG	1800
Db	1741	CAGCCATTGTGAGAGCCAGAGGGGGAACCACTTTTGGGAACCCCCAACCCCTCCATGGATGG	1800
Qy	1801	AGAACTGCTCTGAGGCGAGAGGATCTACGCCAGCAGGTGAGGGTGTGTGAGGGGGTGG	1860
Db	1801	AGAACTGCTCTGAGGCGAGAGGATCTACGCCAGCAGGTGAGGGTGTGTGAGGGGGTGG	1860
Qy	1861	CGGCTTTCAGCCCTCTGGCTTGGCCTTTCCTTTCACAGCTGTAATAATCCCTCCCAATCT	1920
Db	1861	CGGCTTTCAGCCCTCTGGCTTGGCCTTTCCTTTCACAGCTGTAATAATCCCTCCCAATCT	1920
Qy	1921	TCCTTTCCTCTCTTCCCTTTCCTCTCTCCCCCTCGGTGAATGATGGCTGTCTTCTAAAA	1980
Db	1921	TCCTTTCCTCTCTTCCCTTTCCTCTCTCCCCCTCGGTGAATGATGGCTGTCTTCTAAAA	1980
Qy	1981	CAAAATCAACCAAACTCAGCAGTGTGATCTATAGCAGGATGGCCCAAGTACCTGGCTCCA	2040
Db	1981	CAAAATCAACCAAACTCAGCAGTGTGATCTATAGCAGGATGGCCCAAGTACCTGGCTCCA	2040
Qy	2041	CTGATCACCTCTCTCTGTGACCATCAACCAACGGGTGCTCTTGGCCTGGCTTCCCTTG	2100
Db	2041	CTGATCACCTCTCTCTGTGACCATCAACCAACGGGTGCTCTTGGCCTGGCTTCCCTTG	2100
Qy	2101	GCCTTCTCAGCTTCACTTGATGCTGGGCTCTTCTTGTCTCATGTCTCAAGCTGTGGAC	2160
Db	2101	GCCTTCTCAGCTTCACTTGATGCTGGGCTCTTCTTGTCTCATGTCTCAAGCTGTGGAC	2160
Qy	2161	CAGAGACCTGGACTTTGTCTGCTTAAAGGAAATAGGGAAGTAAAGACAGTGAAGGGT	2220
Db	2161	CAGAGACCTGGACTTTTGTCTGCTTAAAGGAAATAGGGAAGTAAAGACAGTGAAGGGT	2220

```
Qy 2221 GGAGGTTGATCAGGCGACAGTGGACAGGAGACCTCACAGAGAAGCGCTGGAAGTGA 2280
Db 2221 GGAGGTTGATCAGGCGACAGTGGACAGGAGACCTCACAGAGAAGCGCTGGAAGTGA 2280
Qy 2281 TTTCCTGCTGACTCATGATAGGATACAAAATGTTCCATGTACCAATTAATCTTGACA 2340
Db 2281 TTTCCTGCTGACTCATGATAGGATACAAAATGTTCCATGTACCAATTAATCTTGACA 2340
Qy 2341 TATGCAATGCATAAGACTTCTATTAAAATAAGCTTTTGGAAAGAGATTAAAAAATAAAA 2400
Db 2341 TATGCAATGCATAAGACTTCTATTAAAATAAGCTTTTGGAAAGAGATTAAAAAATAAAA 2400
Qy 2401 AAAGGCGCGCTCTAGAGGATCAAGCTTACGTPACGCGTGCAATGCGACGTCATAGCTC 2460
Db 2401 AAAGGCGCGCTCTAGAGGATCAAGCTTACGTPACGCGTGCAATGCGACGTCATAGCTC 2460
Qy 2461 TTCTATAGTGTCAACCTAAATTAATT 2486
Db 2461 TTCTATAGTGTCAACCTAAATTAATT 2486
```

RESULT 2

```
US-10-664-667-4
; Sequence 4, Application US/10664667
; Publication No. US20040058377A1
; GENERAL INFORMATION:
; APPLICANT: Gu, Wei
; TITLE OF INVENTION: NOVEL G-PROTEIN COUPLED RECEPTORS AND USES THEREFOR
; FILE REFERENCE: MN1-080CP
; CURRENT APPLICATION NUMBER: US/10/664,667
; PRIOR FILING DATE: 2003-09-18
; PRIOR APPLICATION NUMBER: US/09/851,595
; PRIOR FILING DATE: 2000-05-08
; PRIOR APPLICATION NUMBER: 09/556,588
; PRIOR FILING DATE: 2000-05-08
; PRIOR APPLICATION NUMBER: 60/132,896
; PRIOR FILING DATE: 1999-05-06
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 4
; LENGTH: 2486
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (2)..(1900)
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (172)
; OTHER INFORMATION: n = any nucleotide
```

US-10-664-667-4

```
Query Match 100.0%; Score 2485; DB 17; Length 2486;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2486; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TAATACGACTCACTATAGGAAGCTGTACGCTGCAGGTACCGGTCCGGAATCCCGG 60
Db 1 TAATACGACTCACTATAGGAAGCTGTGTACGCTGCAGGTACCGGTCCGGAATTCCTCGG 60
Qy 61 GTCGACCCACCGCTCCGTGGAGCGGAGCCAGGCTCTGAGCCCTGCGGCTCATCCAGCCTC 120
Db 61 GTCGACCCACCGCTCCGTGGAGCGGAGCCAGGCTCTGAGCCTGCGGCTCATCCAGCCTC 120
Qy 121 TCTTGTGCTTCCCTAGCGGCTCCAAACAACCGCATCTGGAAATTTGGAGCTNGACACCTT 180
Db 121 TCTTGTGCTTCCCTAGCGGCTCCAAACAACCGCATCTGGAAATTTGGAGCTNGACACCTT 180
Qy 181 CAGCCAGCTGAGCTCCCTGCAAGCCCTGGATCTTAGCTTGGAAACGCCATCCCGTCCATCCA 240
Db 181 CAGCCAGCTGAGCTCCCTGCAAGCCCTGGATCTTAGCTTGGAAACGCCATCCCGTCCATCCA 240
```

```
Qy 241 CCCTGAGGCTTTCTCCACCCCTGCATCCCTGGTCAAGCTGCACTGACAGACAACAGCT 300
Db 241 CCCTGAGGCTTTCTCCACCCCTGCATCCCTGGTCAAGCTGCACTGACAGACAACAGCT 300
Qy 301 GACCACATGCGCTGGCTGCACTTGGGGCTTGATGCATCTGAAGCTCAAGGGAACCT 360
Db 301 GACCACATGCGCTGGCTGCACTTGGGGCTTGATGCATCTGAAGCTCAAGGGAACCT 360
Qy 361 TGGCTCTCTCCAGAGGCTTTCTCAAGGACAGTTTCCCAAAACTGAGGATCTCTGAGGTC 420
Db 361 TGGCTCTCTCCAGAGGCTTTCTCAAGGACAGTTTCCCAAAACTGAGGATCTCTGAGGTC 420
Qy 421 TTATGCTTACCAAGTCTGCTCCCTATGGAATGTGCGCAAGCTTCTTCAAGGCTCTGGGCA 480
Db 421 TTATGCTTACCAAGTCTGCTCCCTATGGAATGTGCGCAAGCTTCTTCAAGGCTCTGGGCA 480
Qy 481 GTGGAGGCTGAAGACCTTCACTTGTATGATGAGGAGTCTTCAAAAAGGCGCCCTGGGCT 540
Db 481 GTGGAGGCTGAAGACCTTCACTTGTATGATGAGGAGTCTTCAAAAAGGCGCCCTGGGCT 540
Qy 541 CTTTGCAGACAAGCAGAGAACCACTATGACAGGACCTGGATGAGCTCCAGCTGGAGAT 600
Db 541 CTTTGCAGACAAGCAGAGAACCACTATGACAGGACCTGGATGAGCTCCAGCTGGAGAT 600
Qy 601 GGAGGACTCAAAAGCCACACCCAGTGTCCAGTGTAGCCCTACTCCAGGCGCCCTTCAAGCC 660
Db 601 GGAGGACTCAAAAGCCACACCCAGTGTCCAGTGTAGCCCTACTCCAGGCGCCCTTCAAGCC 660
Qy 661 CTGTGAGTACCTTTTGAAGCTGGGCGATCCGCTGGCGGTGGGGCATCGTGTGCT 720
Db 661 CTGTGAGTACCTTTTGAAGCTGGGCGATCCGCTGGCGGTGGGGCATCGTGTGCT 720
Qy 721 CTCGCTGCTCTGCAATGGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 780
Db 721 CTCGCTGCTCTGCAATGGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 780
Qy 781 GCCCCCGGTCAAAGTTTGTGTGAGTGTGATGCGATGTCAGGCGCCCAACCTTGAATGCGATTC 840
Db 781 GCCCCCGGTCAAAGTTTGTGTGAGTGTGATGCGATGTCAGGCGCCCAACCTTGAATGCGATTC 840
Qy 841 CTGTGGGCTTCTAGCCTCAGTGCATGCGCTGACCTTTGGTCACTTCTGATGAGGAGC 900
Db 841 CTGTGGGCTTCTAGCCTCAGTGCATGCGCTGACCTTTGGTCACTTCTGATGAGGAGC 900
Qy 901 CCGCTGGGAGACGGGCTAGGCTCCGCGGCACCTGGCTTCTGCGCAGTACTTGGGTCGGA 960
Db 901 CCGCTGGGAGACGGGCTAGGCTCCGCGGCACCTGGCTTCTGCGCAGTACTTGGGTCGGA 960
Qy 961 GGCAATCGGTGCTGCTGCTCACTCTGGCGGCAAGTGCAGTGCAGCGTCTCCGTCTCTGTGT 1020
Db 961 GGCAATCGGTGCTGCTGCTCACTCTGGCGGCAAGTGCAGTGCAGCGTCTCCGTCTCTGTGT 1020
Qy 1021 CCGGCGCTATGGGAAGTCCCGCTCCCTGGGAGAGGTTGAGCAGGCGGTCTTAGGCTGCT 1080
Db 1021 CCGGCGCTATGGGAAGTCCCGCTCCCTGGGAGAGGTTGAGCAGGCGGTCTTAGGCTGCT 1080
Qy 1081 GGCACTGGAGGCTGGCGCGGCACCTGCGCTGCGCTCAGTGGGAGAAATACGGGCGCTC 1140
Db 1081 GGCACTGGAGGCTGGCGCGGCACCTGCGCTGCGCTCAGTGGGAGAAATACGGGCGCTC 1140
Qy 1141 CCCACTCTGCTGCCCTACGCGCCACCTGAGGGGTGAGCAGCAGCAGCCCTGGGCTTCAACGT 1200
Db 1141 CCCACTCTGCTGCCCTACGCGCCACCTGAGGGGTGAGCAGCAGCAGCCCTGGGCTTCAACGT 1200
Qy 1201 GGCCCTGCTGATGATGAATCTCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1260
Db 1201 GGCCCTGCTGATGATGAATCTCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1260
Qy 1261 GTACTGTGACCTGCGGGGCGGAGCTTTGAGGCGGTGTGGGACTGCGGCATGCTGAGGCA 1320
Db 1261 GTACTGTGACCTGCGGGGCGGAGCTTTGAGGCGGTGTGGGACTGCGGCATGCTGAGGCA 1320
Qy 1321 CGTGGCTGGCTCATCTTGCAGAGCGGCTCTCTACTGTCCCGTGGGCTTCTCTCAGCTT 1380
```

[illegible]

Db	2401	AAAGGGGGGGCGCTCTAGAGGATCCAAAGCTTACGTACGCGTGCATCGCAGCTCANTAGCTC	2486
Qy	2461	TTCTATAGTGTACCTAAATTTCAATT	2486
Db	2461	TTCTATAGTGTACCTAAATTTCAATT	2486
RESULT 3			
US-09-851-595-7			
; Sequence 7, Application US/09851595			
; Publication No. US20030166047A1			
; GENERAL INFORMATION:			
; APPLICANT: Gu, Wei			
; TITLE OF INVENTION: NOVEL G-PROTEIN COUPLED RECEPTORS AND USES THEREFOR			
; FILE REFERENCE: MWI-080CP			
; CURRENT APPLICATION NUMBER: US/09/851,595			
; CURRENT FILING DATE: 2000-05-08			
; PRIOR APPLICATION NUMBER: 09/556,588			
; PRIOR FILING DATE: 2000-05-08			
; PRIOR APPLICATION NUMBER: 60/132,896			
; PRIOR FILING DATE: 1999-05-06			
; NUMBER OF SEQ ID NOS: 12			
; SOFTWARE: PatentIn Ver. 2.0			
; SEQ ID NO 7			
; LENGTH: 2711			
; TYPE: DNA			
; ORGANISM: Homo sapiens			
; FEATURE:			
; NAME/KEY: CDS			
; LOCATION: (1)..(2208)			
US-09-851-595-7			
Query Match 90.8%; Score 2257.4; DB 10; Length 2711;			
Best Local Similarity 99.9%; Pred. No. 0;			
Matches 2269; Conservative 0; Mismatches 1; Indels 1; Gaps 1;			
Qy	133	AGCGGGCTCCAAACACACGCGCATCTGGGAAATTTGGAGCTNGACACCTTCAGCCAGCTGAG	192
Db	442	ATCGGCTTCCAAACACACGCGCATCTGGGAAATTTGGAGCT-GACACCTTCAGCCAGCTGAG	500
Qy	193	CTCCCTGCAAGCCCTGGATCTTAGCTGGAAAGCGCATCCGGTTCATCCACCTGAGGCGCTT	252
Db	501	CTCCCTGCAAGCCCTGGATCTTAGCTGGAAAGCGCATCCGGTTCATCCACCTGAGGCGCTT	560
Qy	253	CTCCACCTTGCACTCCCTGGTCAAGCTGCACTGACAGACACCACTGACACCACTGCC	312
Db	561	CTCCACCTTGCACTCCCTGGTCAAGCTGCACTGACAGACCACTGACACCACTGCC	620
Qy	313	CTGGCTGGACTTGGGGGCTTGATGATCTGAAGCTCAAAGGGAACCTTGCTCTCTCCCA	372
Db	621	CTGGCTGGACTTGGGGGCTTGATGATCTGAAGCTCAAAGGGAACCTTGCTCTCTCCCA	680
Qy	373	GGCCTTCTCAAGGACAGTTTCCCAAACTGAGGATCTCTGGAGTGCTTATGCCTACCA	432
Db	681	GGCCTTCTCAAGGACAGTTTCCCAAACTGAGGATCTCTGGAGTGCTTATGCCTACCA	740
Qy	433	GTGCTGTCCCTATGGATGTGTCGACAGCTTCTTCAAGGCTCTGGGAGTGGGAGGCTGA	492
Db	741	GTGCTGTCCCTATGGGATGTGTCGACAGCTTCTTCAAGGCTCTGGGAGTGGGAGGCTGA	800
Qy	493	AGACCTTACCTTGATGATGAGGAGTCTTCAAAAAGGCCCTTGGGCTCTCTTGGCCAGACA	552
Db	801	AGACCTTACCTTGATGATGAGGAGTCTTCAAAAAGGCCCTTGGGCTCTCTTGGCCAGACA	860
Qy	553	AGCAGAGAACCACTATGACAGGACCTTGATGAGCTTCCAGCTTGGAGATGGAGGACTCAAA	612
Db	861	AGCAGAGAACCACTATGACAGGACCTTGATGAGCTTCCAGCTTGGAGATGGAGGACTCAAA	920
Qy	613	GCCACACCCAGTGTCCAGTGTAGCCCTACTCCAGGCCCTTTCAGGCCCTGTGAGTACCT	672
Db	921	GCCACACCCAGTGTCCAGTGTAGCCCTACTCCAGGCCCTTTCAGGCCCTGTGAGTACCT	980
Qy	673	CTTTGAAAGCTGGGGCATCCGGCTTGCGGCGCATCGTGTGTGTGCTCTCCGTCGCTCTG	732

Db	981	CTTTGAAAGCTGGGGCAATCCGCCTGGCCGCTGGCCGCAATCGTGTGCTCTCCCGTGTCTG	1040
Qy	733	CAATGGACTGGTGTCTGTGACCGGTGTTGCGTGGCGGCGCTGCCCGCTGCCCGGTCAA	792
Db	1041	CAATGGACTGGTGTCTGTGACCGGTGTTGCGTGGCGGCGCTGCCCGCTGCCCGGTCAA	1100
Qy	793	GTTTGTGGTAGTGCGAATTGACGGCCAAACACTTGACTGGCAATTCCTGTGGCCTTCT	852
Db	1101	GTTTGTGGTAGTGCGAATTGACGGCCAAACACTTGACTGGCAATTCCTGTGGCCTTCT	1160
Qy	853	AGCCTCAGTCGATGCCCTGACCTTTGCTGAGTTCCTGAGTACGGAGCCCGCTGGGAGAC	912
Db	1161	AGCCTCAGTCGATGCCCTGACCTTTGCTGAGTTCCTGAGTACGGAGCCCGCTGGGAGAC	1220
Qy	913	GGGCTTAGGCTGCCGGGCCACTGGCTTCCTGGGAGTACTTGGGTGCGAGGACATCGGTGCT	972
Db	1221	GGGCTTAGGCTGCCGGGCCACTGGCTTCCTGGGAGTACTTGGGTGCGAGGACATCGGTGCT	1280
Qy	973	GCTGCTCACTCTGGCCGCGAGTGACGAGGCTTCGCTCTCTGTGTGCGGCGCTATGG	1032
Db	1281	GCTGCTCACTCTGGCCGCGAGTGACGAGGCTTCGCTCTCTGTGTGCGGCGCTATGG	1340
Qy	1033	GAAGTCCCCCTCCCTGGGACGTTTCAGACAGGGGTCTTAGGCTGCTGGCATGGCAGG	1092
Db	1341	GAAGTCCCCCTCCCTGGGACGTTTCAGACAGGGGTCTTAGGCTGCTGGCATGGCAGG	1400
Qy	1093	GCTGGCCCGCGCACTGCCCTTGGGCTCAGTGGGAGAAACAGGGGCTTCCGCCACTCTGCGCT	1152
Db	1401	GCTGGCCCGCGCACTGCCCTTGGGCTCAGTGGGAGAAACAGGGGCTTCCGCCACTCTGCGCT	1460
Qy	1153	GCCTTACGCGCCACTGAGGGTCAGCAGAGCCCTGGGGTTACCGTGGCCCTGCTGTGAT	1212
Db	1461	GCCTTACGCGCCACTGAGGGTCAGCAGAGCCCTGGGGTTACCGTGGCCCTGCTGTGAT	1520
Qy	1213	GATGAACCTCCTTCGTGTTCTGTCGTGGCGCGTGCCCTACATCAAATGTCATGTGACCT	1272
Db	1521	GATGAACCTCCTTCGTGTTCTGTCGTGGCGCGTGCCCTACATCAAATGTCATGTGACCT	1580
Qy	1273	GC CGCGGGCGACTTTGAGGCGGTGGGACTGGCCATGGTAGGCACTGTGGCCCTGGCT	1332
Db	1581	GC CGCGGGCGACTTTGAGGCGGTGGGACTGGCCATGGTAGGCACTGTGGCCCTGGCT	1640
Qy	1333	CATCTTCGACAGCGGCTCTCTACTGTCCCGTGGCCCTCCTCAGCTTGCCTCCATGCT	1392
Db	1641	CATCTTCGACAGCGGCTCTCTACTGTCCCGTGGCCCTCCTCAGCTTGCCTCCATGCT	1700
Qy	1393	GGGCTCTTCCCTGTACGCCCGAGGCGCTCAAGTCTGCTCGTGGTGTGTGCTGCCCT	1452
Db	1701	GGGCTCTTCCCTGTACGCCCGAGGCGCTCAAGTCTGCTCGTGGTGTGTGCTGCCCT	1760
Qy	1453	GCCTGCTGCTCAACCCACTGCTGTACCTGCTTTCACCCCGCCACTTCGGGATGACCT	1512
Db	1761	GCCTGCTGCTCAACCCACTGCTGTACCTGCTTTCACCCCGCCACTTCGGGATGACCT	1820
Qy	1513	TCGGCGGCTTCGGCCCGCGCAGGGGACTCAGGGCCCTTAGCCCTATGCTCGGCCCGGGA	1572
Db	1821	TCGGCGGCTTCGGCCCGCGCAGGGGACTCAGGGCCCTTAGCCCTATGCTCGGCCCGGGA	1880
Qy	1573	GCTGGAGAAGCTCTGTGATTTACCAAGGCCCTGGTAGCCTTCTCTGATGTGATCT	1632
Db	1881	GCTGGAGAAGCTCTGTGATTTACCAAGGCCCTGGTAGCCTTCTCTGATGTGATCT	1940
Qy	1633	CATTCTGGAAGCTTCTCAAGCTGGCGGCCCTGGCTGGAGACCTATGGCTTCCCTC	1692
Db	1941	CATTCTGGAAGCTTCTCAAGCTGGCGGCCCTGGCTGGAGACCTATGGCTTCCCTC	2000
Qy	1693	AGTGACCTCATCTCTGTGTCAGACGACAGGGGCCCCCAGGCTGGAGGGCAGCAATGTGT	1752
Db	2001	AGTGACCTCATCTCTGTGTCAGACGACAGGGGCCCCCAGGCTGGAGGGCAGCAATGTGT	2060
Qy	1753	AGAGCCAGAGGGGAACCACTTTGGGAACCCCCAAACCTCTCCATGGATGGAGAACTGTGCT	1812

Db	2061	AGAGCCAGAGGGGAAACCACTTTTGGGAACCCCAACCCCTCCATGGATGGAGAACTGTCTGCT	2120
Qy	1813	GAGGGCAGAGGGATCTACGCCAGCAGAGTGGAGGCTTGTTCAGGGGGTGGCGCTTTTCAGCC	1872
Db	2121	GAGGGCAGAGGGATCTACGCCAGCAGAGTGGAGGCTTGTTCAGGGGGTGGCGCTTTTCAGCC	2180
Qy	1873	CTCTGGCTTTGGCCTTTGGCTTCAACAGTGTAATAATCCCTCCCAATTTCTTCTTTCCCTC	1932
Db	2181	CTCTGGCTTTGGCCTTTGGCTTCAACAGTGTAATAATCCCTCCCAATTTCTTCTTTCCCTC	2240
Qy	1933	TCCTCCCTTTCTCTCTCCCTCGGTGAATGATGGCTGCTCTTAAACAAATACAAACCA	1992
Db	2241	TCCTCCCTTTCTCTCTCCCTCGGTGAATGATGGCTGCTCTTAAACAAATACAAACCA	2300
Qy	1993	AAACTCAGCAGTGTGATCTATAGCAGGATGGCCAGTACTGGCTCCACTGATCACTCT	2052
Db	2301	AAACTCAGCAGTGTGATCTATAGCAGGATGGCCAGTACTGGCTCCACTGATCACTCT	2360
Qy	2053	CTCCGTGACCATCACCAACGGGTGCTCTTGGCCTGGCTTTCCCTTGGGCTTCTTCAGC	2112
Db	2361	CTCCGTGACCATCACCAACGGGTGCTCTTGGCCTGGCTTTCCCTTGGGCTTCTTCAGC	2420
Qy	2113	TTCACTTTGATCTAGGGCCTCTTCTTGTCTCATGTCTGAAGCTGTGGACACAGACCTGGA	2172
Db	2421	TTCACTTTGATCTAGGGCCTCTTCTTGTCTCATGTCTGAAGCTGTGGACACAGACCTGGA	2480
Qy	2173	CTTTTGTCTCTTAAGGGAATGAGGGAAGTAAAGACAGTGAAGGGGTGGAGGTTGATC	2232
Db	2481	CTTTTGTCTCTTAAGGGAATGAGGGAAGTAAAGACAGTGAAGGGGTGGAGGTTGATC	2540
Qy	2233	AGGCGACAGTGGACAGGGAGACCTCACAGAGAAGGCCTGGAGGTTGATTTCCCGTGTGA	2292
Db	2541	AGGCGACAGTGGACAGGGAGACCTCACAGAGAAGGCCTGGAGGTTGATTTCCCGTGTGA	2600
Qy	2293	CTCATGATAGGATACAAAATGTGTTCCATGTACCATTAACTTTGACATATGCCATGCGAT	2352
Db	2601	CTCATGATAGGATACAAAATGTGTTCCATGTACCATTAACTTTGACATATGCCATGCGAT	2660
Qy	2353	AAAGACTTCCTATTTAAATTAAGCTTTTGGAAAGAGATTAAAAAATAAAAAA	2403
Db	2661	AAAGACTTCCTATTTAAATTAAGCTTTTGGAAAGAGATTAAAAAATAAAAAA	2711

DECIU.T A

```

RESUDI 4
US-10-664-667-7
; Sequence 7, Application US/10664667
; Publication No. US20040058377A1
; GENERAL INFORMATION:
; APPLICANT: Gu, Wei
; TITLE OF INVENTION: NOVEL G-PROTEIN COUPLED RECEPTORS AND USES THEREFOR
; FILE REFERENCE: WNI-080CP
; CURRENT APPLICATION NUMBER: US/10/664,667
; CURRENT FILING DATE: 2003-09-18
; PRIOR APPLICATION NUMBER: US/09/851,595
; PRIOR FILING DATE: 2000-05-08
; PRIOR APPLICATION NUMBER: 09/556,588
; PRIOR FILING DATE: 2000-05-08
; PRIOR APPLICATION NUMBER: 60/132,896
; PRIOR FILING DATE: 1999-05-06
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: PatentIn ver. 2.0
; SEQ ID NO 7
; LENGTH: 2711
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(2208)
US-10-664-667-7

```

Query Match 90.8%; Score 2257.4; DB 17; Length 2711;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 2269; Conservative 0; Mismatches 1; Indels 1;

Qy	133	AGGSCCTCCAAACACAAACCGCATCTGGGAAATTTGGAGCTNGACACCTTTCAAGCCAGCTGAG	192
Db	442	ATCGGCCTCCAAACACAAACCGCATCTGGGAAATTTGGAGCT -GACACCTTTCAAGCCAGCTGAG	500
Qy	193	CTCCCTGAAGCCCTTGATCTTTAGCTGGAAAGCCATCCGGTCCCATCCACCCCTGAGGCCCTT	252
Db	501	CTCCCTGAAGCCCTTGATCTTTAGCTGGAAAGCCATCCGGTCCCATCCACCCCTGAGGCCCTT	560
Qy	253	CTCACCCCTGCATCTCCCTGGTCAAGCTGGACCTGACAGACAAACAGCTGACCAACATGACC	312
Db	561	CTCACCCCTGCATCTCCCTGGTCAAGCTGGACCTGACAGACAAACAGCTGACCAACATGACC	620
Qy	313	CTTGCTGGAATCTGGGGCTTTGATGCACTCTGAAGCTCAAGGGAACCTTGCTCTCTCCCA	372
Db	621	CTTGCTGGAATCTGGGGCTTTGATGCACTCTGAAGCTCAAGGGAACCTTGCTCTCTCCCA	680
Qy	373	GGCCTTCTCCAAAGACAGTTTCCCAAAACTGAGGATCTCGAGGTGCTTATGCTTACCA	432
Db	681	GGCCTTCTCCAAAGACAGTTTCCCAAAACTGAGGATCTCGAGGTGCTTATGCTTACCA	740
Qy	433	GTGCTGTCCCTATGGGATGTGTCAGCTTTCTCAAGGCCCTCTGGGCAGTGGGAGGTGA	492
Db	741	GTGCTGTCCCTATGGGATGTGTCAGCTTTCTCAAGGCCCTCTGGGCAGTGGGAGGTGA	800
Qy	493	AGACCTTACCTTGATGATGAGGAGTCTTCAAAAGGCCCTTGGGCCCTCTTGGCCAGACA	552
Db	801	AGACCTTACCTTGATGATGAGGAGTCTTCAAAAGGCCCTTGGGCCCTCTTGGCCAGACA	860
Qy	553	AGCAGAGAAACCATATGACAGGACCTTGGATGAGCTCCAGCTGGAGATGGAGGACTCAAA	612
Db	861	AGCAGAGAAACCATATGACAGGACCTTGGATGAGCTCCAGCTGGAGATGGAGGACTCAAA	920
Qy	613	GCCACACCCCAAGTGCAGTGTAGCCCTTATCCAGGCCCTTCAAGGCCCTGTGAGTACCT	672
Db	921	GCCACACCCCAAGTGCAGTGTAGCCCTTATCCAGGCCCTTCAAGGCCCTGTGAGTACCT	980
Qy	673	CTTTGAAAGCTGGGGCATCCGCTTGGCGTGGCGTGGCCATCGTGTGCTCTCCGTGCTCTG	732
Db	981	CTTTGAAAGCTGGGGCATCCGCTTGGCGTGGCGTGGCCATCGTGTGCTCTCCGTGCTCTG	1040
Qy	733	CAATGGAGCTGGTGTGCTGACCGTGTTCGCTGGGGGCTTCCGCCCTTCCGCCCTTCA	792
Db	1041	CAATGGAGCTGGTGTGCTGACCGTGTTCGCTGGGGGCTTCCGCCCTTCCGCCCTTCA	1100
Qy	793	GTTTGTGTAGGTGGGATGTGAGGGCCAAACCTTGAATGGGATTTCTGTGGCCCTTCT	852
Db	1101	GTTTGTGTAGGTGGGATGTGAGGGCCAAACACCTTGAATGGGATTTCTGTGGCCCTTCT	1160
Qy	853	AGCCTCAGTGCATGCCCTGACCTTTGGTTCAGTTCTCTGAGTACGAGGCCCTGGGAGAC	912
Db	1161	AGCCTCAGTGCATGCCCTGACCTTTGGTTCAGTTCTCTGAGTACGAGGCCCTGGGAGAC	1220
Qy	913	GGGGCTAGGCTGCGGGCCACTGGCTTCTGGCAGTACTTGGGTGGAGGCATCGGTGCT	972
Db	1221	GGGGCTAGGCTGCGGGCCACTGGCTTCTGGCAGTACTTGGGTGGAGGCATCGGTGCT	1280
Qy	973	GCTGCTCACTGTGCCCGAGTGCAGTGCAGCGTCTCCGTCTCTGTGTGCCGGCCTATGG	1032
Db	1281	GCTGCTCACTGTGCCCGAGTGCAGTGCAGCGTCTCCGTCTCTGTGTGCCGGCCTATGG	1340
Qy	1033	GAAGTCCCTCTCCCTGGGACGCTTCGACGAGGGCTCTAGGCTGCTTGGCACTGGCAGG	1092
Db	1341	GAAGTCCCTCTCCCTGGGACGCTTCGACGAGGGCTCTAGGCTGCTTGGCACTGGCAGG	1400
Qy	1093	GCTGGCCGCGCAGCTGCCCTTGGGCTCAGTGGGAGAAATACGGGGCTCTCCCACTCTGCCT	1152
Db	1401	GCTGGCCGCGCAGCTGCCCTTGGGCTCAGTGGGAGAAATACGGGGCTCTCCCACTCTGCCT	1460
Qy	1153	GCCCTACGCGCCACTGAGGGTCAAGCAGAGCCCTGGGCTTCAACCTGGGCCCTGTGTAT	1212
Db	1461	GCCCTACGCGCCACTGAGGGTCAAGCAGAGCCCTGGGCTTCAACCTGGGCCCTGTGTAT	1520

Qy	1213	GATGAACCTCTCTCTGTTTCTGTTCTGGTCGGCGGTGCTCATCAAACTGTACTGTGACCT	1272
Db	1521	GATGAACCTCTCTCTGTTTCTGTTCTGGTCGGCGGTGCTCATCAAACTGTACTGTGACCT	1580
Qy	1273	GCOCGGGGGCACTTTGAGGCCGTTGCGGACTGCGCCATGGTGAAGGCACTGGGCTTGCT	1332
Db	1581	GCOCGGGGGCACTTTGAGGCCGTTGCGGACTGCGCCATGGTGAAGGCACTGGGCTTGCT	1640
Qy	1333	CATCTTCGACAGACGGGCTCTCTACTGTCCGGTGGCCCTCTCTCAGCTTGGCTTCATGCT	1392
Db	1641	CATCTTCGACAGACGGGCTCTCTACTGTCCGGTGGCCCTCTCTCAGCTTGGCTTCATGCT	1700
Qy	1393	GGGCTCTCTCCCTGTCAAGCCCGAGGCCGTTCAAGTCTGTCTCTGTGTGTGTGTCGCCCT	1452
Db	1701	GGGCTCTCTCTCCCTGTCAAGCCCGAGGCCGTTCAAGTCTGTCTCTGTGTGTGTGTCGCCCT	1760
Qy	1453	GCCTGCTGCTCTCAACCCACTGTGTACTGCTCTTCAACCCCCACCTTCCGGGATGACCT	1512
Db	1761	GCCTGCTGCTCTCAACCCACTGTGTACTGCTCTTCAACCCCCACCTTCCGGGATGACCT	1820
Qy	1513	TCGGCGGCTTCGGCCCGCGCAGGGGACTCAGGGCCCTAGCCTATGCTGCGGCCGGGA	1572
Db	1821	TCGGCGGCTTCGGCCCGCGCAGGGGACTCAGGGCCCTAGCCTATGCTGCGGCCGGGA	1880
Qy	1573	GCTGGAGAAAGAGCTCCCTGTGATTTACCCAGGCCCTGGTAGCCTTCTCTGATGTGGATCT	1632
Db	1881	GCTGGAGAAAGAGCTCCCTGTGATTTACCCAGGCCCTGGTAGCCTTCTCTGATGTGGATCT	1940
Qy	1633	CATTCTGGAAGCTTCTGAAAGCTGGGGGCCCTTGGGCTGGAGACCTATGGCTTCCCTCTC	1692
Db	1941	CATTCTGGAAGCTTCTGAAAGCTGGGGGCCCTTGGGCTGGAGACCTATGGCTTCCCTCTC	2000
Qy	1693	AGTGAACCTCATCTCTGTGACAGACGAGGGGCCCTCAGGCTGGAGGAGACCATTTGTGT	1752
Db	2001	AGTGAACCTCATCTCTGTGACAGACGAGGGGCCCTCAGGCTGGAGGAGACCATTTGTGT	2060
Qy	1753	AGAGCCAGAGGGGAAACCATTTTGGGAAACCCCAACCTCCATCGATGAGAGAACTGTGCT	1812
Db	2061	AGAGCCAGAGGGGAAACCATTTTGGGAAACCCCAACCTCCATCGATGAGAGAACTGTGCT	2120
Qy	1813	GAGGCGAGAGGGATCTACGCCAGAGTGAAGCTTGTCAAGGGGTGGCGGCTTTCAGCC	1872
Db	2121	GAGGCGAGAGGGATCTACGCCAGAGTGAAGCTTGTCAAGGGGTGGCGGCTTTCAGCC	2180
Qy	1873	CTCTGGCTGGCCCTTTGCTTCAACGTGTAAATATCCCTCCCACTTCTTCTTCTTCCCTCTC	1932
Db	2181	CTCTGGCTGGCCCTTTGCTTCAACGTGTAAATATCCCTCCCACTTCTTCTTCTTCCCTCTC	2240
Qy	1933	TCCTTCCCTTTCCTCTCTCCCTCCGTTGATGATGGCTGCTTCTTAAACAAATACAAACCA	1992
Db	2241	TCCTTCCCTTTCCTCTCTCCCTCCGTTGATGATGGCTGCTTCTTAAACAAATACAAACCA	2300
Qy	1993	AAACTCAGCAGTGTGATCTATAGCAGGATGGCCCACTGACCTGGCTCCACTGATCACCTCT	2052
Db	2301	AAACTCAGCAGTGTGATCTATAGCAGGATGGCCCACTGACCTGGCTCCACTGATCACCTCT	2360
Qy	2053	CTCTGTGACCATCAACAGGGTGCCTCTTGGCGCTGGCTTTCCCTTGGCGCTTCTCAGC	2112
Db	2361	CTCTGTGACCATCAACAGGGTGCCTCTTGGCGCTGGCTTTCCCTTGGCGCTTCTCAGC	2420
Qy	2113	TTCACTTGTACTTGGGCTCTTCTTGTATCTGTCTGTAAGCTGTGGACACAGAGACCTTGA	2172
Db	2421	TTCACTTGTACTTGGGCTCTTCTTGTATCTGTCTGTAAGCTGTGGACACAGAGACCTTGA	2480
Qy	2173	CTTTTGTCTGCTTAAAGGAAATGAGGAAAGTAAAGACAGTGAAGGGGTGGAGGGTTGATC	2232
Db	2481	CTTTTGTCTGCTTAAAGGAAATGAGGAAAGTAAAGACAGTGAAGGGGTGGAGGGTTGATC	2540
Qy	2233	AGGGCAGTGTGACAGGGAGACCTCAACAGAAAGCCCTGGAAAGGTGATTTCCCGTGTGA	2292
Db	2541	AGGGCAGTGTGACAGGGAGACCTCAACAGAAAGCCCTGGAAAGGTGATTTCCCGTGTGA	2600
Qy	2293	CTCATGATAGGATACAAATGTGTTCCATGTACCATTAATCTTGACATATGCATGCAT	2352

QY 1693 AGTGACCTCACTCTCTGTGACAGCAGGGGCCCCCAGGCTGGAGGCGAGCATTGTGT 1752
 DB |||||
 QY 2277 AGTGACCTCATCTCTGTGACAGCAGGGGCCCCCAGGCTGGAGGCGAGCATTGTGT 2336
 DB |||||
 QY 1753 AGAGCCAGAGGGGAACCACTTTGGGAACCCCAACCCCTCCATGGATGAGAACTGCTGCT 1812
 DB |||||
 QY 2337 AGAGCCAGAGGGGAACCACTTTGGGAACCCCAACCCCTCCATGGATGAGAACTGCTGCT 2396
 DB |||||
 QY 1813 GAGGCGAGAGGATCTAGCCAGCAGGTGGAGGCTTGTTCAGGGGTGGGCGCTTCAGCC 1872
 DB |||||
 QY 2397 GAGGCGAGAGGATCTAGCCAGCAGGTGGAGGCTTGTTCAGGGGTGGGCGCTTCAGCC 2456
 DB |||||
 QY 1873 CTCTGGCTTGGCTTGTCTTACAGCTGTAATAATATCCCTCCCACTTCTCTCTCCCTC 1932
 DB |||||
 QY 2457 CTCTGGCTTGGCTTGTCTTACAGCTGTAATAATATCCCTCCCACTTCTCTCTCCCTC 2516
 DB |||||
 QY 1933 TCTTCCCTTCTCTCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCT 1992
 DB |||||
 QY 2517 TCTTCCCTTCTCTCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCT 2576
 DB |||||
 QY 1993 AAACCTCAGCAGTGTGATCTATAGCAGGATGGCCAGTACCTGGCTCCACTGATCACCTCT 2052
 DB |||||
 QY 2577 AAACCTCAGCAGTGTGATCTATAGCAGGATGGCCAGTACCTGGCTCCACTGATCACCTCT 2636
 DB |||||
 QY 2053 CTCTGTGACCATCAACCAAGGCTGCTTCTGGCTGGCTTCTCCCTGGCTTCTCTCAGC 2112
 DB |||||
 QY 2637 CTCTGTGACCATCAACCAAGGCTGCTTCTGGCTGGCTTCTCCCTGGCTTCTCTCAGC 2696
 DB |||||
 QY 2113 TTCACCTTGTATCTGGGCT 2172
 DB |||||
 QY 2697 TTCACCTTGTATCTGGGCT 2756
 DB |||||
 QY 2173 CTTTGTCTGTCTTAAGGGAATAGAGGAAGTAAAGACAGTGAAGGGGTGGAGGTTGATC 2232
 DB |||||
 QY 2757 CTTTGTCTGTCTTAAGGGAATAGAGGAAGTAAAGACAGTGAAGGGGTGGAGGTTGATC 2816
 DB |||||
 QY 2233 AGGCGACAGTGGACAGGAGACCTCACAGAGAAAGGCTCGAAGTGATTTCCCGTGTGA 2292
 DB |||||
 QY 2817 AGGCGACAGTGGACAGGAGACCTCACAGAGAAAGGCTCGAAGTGATTTCCCGTGTGA 2876
 DB |||||
 QY 2293 CTCATGTAGATAGGATACAAATGTGTTCATGTACATTAATTAATTTGACATATGCCATGCAT 2352
 DB |||||
 QY 2877 CTCATGTAGATAGGATACAAATGTGTTCATGTACATTAATTAATTTGACATATGCCATGCAT 2936
 DB |||||
 QY 2353 AAAGACTTCTTAAATAAGCTTTGGAAGAGATTAAAAA 2403
 DB |||||
 QY 2937 AAAGACTTCTTAAATAAGCTTTGGAAGAGATTAAAAA 2987
 DB |||||

RESULT 6

US-10-270-336-4
 ; Sequence 4, Application US/10270336
 ; Publication No. US20030074678A1
 ; GENERAL INFORMATION:
 ; APPLICANT: ZHU, Shiaoqing et al.
 ; TITLE OF INVENTION: ISOLATED HUMAN G-PROTEIN COUPLED
 ; TITLE OF INVENTION: RECEPTORS, NUCLEIC ACID MOLECULES ENCODING HUMAN GPCR
 ; TITLE OF INVENTION: PROTEINS, AND USES THEREOF
 ; FILE REFERENCE: CL001146CON
 ; CURRENT FILING DATE: 2002-10-15
 ; PRIOR FILING DATE: 2001-02-26
 ; NUMBER OF SEQ ID NOS: 7
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 4
 ; LENGTH: 3349
 ; TYPE: DNA
 ; ORGANISM: Human
 US-10-270-336-4

Query Match 90.4%; Score 2247.8; DB 14; Length 3349;
 Best Local Similarity 99.6%; Pred. No. 0;

Matches 2263; Conservative 0; Mismatches 7; Indels 1; Gaps 1;
 QY 133 AGCGGCTTCCAAACAACACCGCATCTGGAAATTTGGAGCTNGACACCTTTCAGCCAGCTGAG 192
 DB |||||
 QY 1060 ATCGGCTTCCAAACAACACCGCATCTGGAAATTTGGAGCT -GACACCTTTCAGCCAGCTGAG 1118
 DB |||||
 QY 193 CTCCCTCAGAGCCCTGTAGCTGGAACGCGCATCCGGTCCATCCACCTCAGGAGGCTT 252
 DB |||||
 QY 1119 CTCCCTCAGAGCCCTGTAGCTGGAACGCGCATCCGGTCCATCCACCCGAGGCGCTT 1178
 DB |||||
 QY 253 CTCCACCTCTGACATCCCTGTGTCAAGCTGGACCTGACAGACAAACAGCTGACACACATGCC 312
 DB |||||
 QY 1179 CTCCACCTCTGACATCCCTGTGTCAAGCTGGACCTGACAGACAAACAGCTGACACACATGCC 1238
 DB |||||
 QY 313 CTGGCTTGGACTTGGGGGCTTGTAGCTGGAAGCTCAAAGGGAACCTTGTCTCTCCCA 372
 DB |||||
 QY 1239 CTGGCTTGGACTTGGGGGCTTGTAGCTGGAAGCTCAAAGGGAACCTTGTCTCTCCCA 1298
 DB |||||
 QY 373 GGCTTCTTCCAAGGACAGTTTCCCAAACTGAGGATCCTTGAGGTGCTTATGCTTACCA 432
 DB |||||
 QY 1299 GGCTTCTTCCAAGGACAGTTTCCCAAACTGAGGATCCTTGAGGTGCTTATGCTTACCA 1358
 DB |||||
 QY 433 GTGCTGTCCCTATGGGATGTGTCCAGCTTCTTCAAGGCTCTCTGGGAGTGGAGGCTGA 492
 DB |||||
 QY 1359 GTGCTGTCCCTATGGGATGTGTCCAGCTTCTTCAAGGCTCTCTGGGAGTGGAGGCTGA 1418
 DB |||||
 QY 493 AGACCTTACCTTGTATGATGAGGATCTTCAAAAGGCTCTGGGCTCTCTTGCAGACA 552
 DB |||||
 QY 1419 AGACCTTACCTTGTATGAGGATCTTCAAAAGGCTCTGGGCTCTCTTGCAGACA 1478
 DB |||||
 QY 553 AGCAGAGAACACATATACACAGGACCTTGGATGAGCTCCAGCTGGAGATGGAGGACTCAAA 612
 DB |||||
 QY 1479 AGCAGAGAACACATATACACAGGACCTTGGATGAGCTCCAGCTGGAGATGGAGGACTCAAA 1538
 DB |||||
 QY 613 GCCACACCCAGTGTCCAGTGTAGCCCTTACTCCAGGCTCTTCAAGGCTCTGTAGTACCT 672
 DB |||||
 QY 1539 GCCACACCCAGTGTCCAGTGTAGCCCTTACTCCAGGCTCTTCAAGGCTCTGTAGTACCT 1598
 DB |||||
 QY 673 CTTTGAAGCTTGGGCTATCCGCTGGGCTGTGGGCTATCGTGTGCTCTCCGCTGCTG 732
 DB |||||
 QY 1599 CTTTGAAGCTTGGGCTATCCGCTGGGCTGTGGGCTATCGTGTGCTCTCCGCTGCTG 1658
 DB |||||
 QY 733 CAATGGACTGTGTCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 792
 DB |||||
 QY 1659 CAATGGACTGTGTCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1718
 DB |||||
 QY 793 GTTTGTGTAGTGTGCGATTTGCGAGCGCAACACCTTGACTGGCATTTCTGTGGCTTCT 852
 DB |||||
 QY 1719 GTTTGTGTAGTGTGCGATTTGCGAGCGCAACACCTTGACTGGCATTTCTGTGGCTTCT 1778
 DB |||||
 QY 853 AGCTCAGTGTGATGCCCTGACCTTTGCTGAGTGTCTGTAGTACGGAGCCGCTGGGAGAC 912
 DB |||||
 QY 1779 AGCTCAGTGTGATGCCCTGACCTTTGCTGAGTGTCTGTAGTACGGAGCCGCTGGGAGAC 1838
 DB |||||
 QY 913 GGGCTAGGCTGCGGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 972
 DB |||||
 QY 1839 GGGCTAGGCTGCGGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1898
 DB |||||
 QY 973 GCTGCTCACTCTGGCGCAGTGCAGTGCAGCGTCTCCGCTCTCTGCTGCTGCTGCTGCTGCTG 1032
 DB |||||
 QY 1899 GCTGCTCACTCTGGCGCAGTGCAGTGCAGCGTCTCCGCTCTCTGCTGCTGCTGCTGCTGCTG 1958
 DB |||||
 QY 1033 GAAGTCCCTCTCCCTGGGAGCGTTCGAGCAGGCTTCTAGGCTGCTGCTGCTGCTGCTGCTG 1092
 DB |||||
 QY 1959 GAAGTCCCTCTCCCTGGGAGCGTTCGAGCAGGCTTCTAGGCTGCTGCTGCTGCTGCTGCTG 2018
 DB |||||
 QY 1093 GCTGGCGCGCAGTGCAGTGCAGCGTCTCCGCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTG 1152
 DB |||||
 QY 2019 GCTGGCGCGCAGTGCAGTGCAGCGTCTCCGCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTG 2078
 DB |||||
 QY 1153 GCCTTAGCGGCCACCTTAGGGTTCAGCCAGCAGCTTGGGCTTCAACCGTGGCCCTTGGTGTAT 1212
 DB |||||
 QY 2079 GCCTTAGCGGCCACCTTAGGGTTCAGCCAGCAGCTTGGGCTTCAACCGTGGCCCTTGGTGTAT 2138
 DB |||||

QY 373 GGCCTTCTCAAGAGACAGTTTCCCAAACTGAGGATCCTGAGGTGCCTTATGCTACCA 432
DB 1302 GGCCTTCTCAAGAGACAGTTTCCCAAACTGAGGATCCTGAGGTGCCTTATGCTACCA 1361
QY 433 GTGCTGCTCCCTATGAGGATGTGTGCGAGCTTCTTCAAGGCCTCTGCGCAGTGGAGGCTGA 492
DB 1362 GTGCTGCTCCCTATGAGGATGTGTGCGAGCTTCTTCAAGGCCTCTGCGCAGTGGAGGCTGA 1421
QY 493 AGACCTTCACTTATGATGATGAGGAGTCTTCAAAAAGGCCCTTGGCCCTCTTGGCAGACA 552
DB 1422 AGACCTTCACTTATGATGATGAGGAGTCTTCAAAAAGGCCCTTGGCCCTCTTGGCAGACA 1481
QY 553 AGCAGAGAACACTATGATACAGGACCTCGATGAGCTCCAGCTGGAGATGGAGGACTCAAA 612
DB 1482 AGCAGAGAACACTATGATACAGGACCTCGATGAGCTCCAGCTGGAGATGGAGGACTCAAA 1541
QY 613 GCCACACCCCAAGTGTCCAGTGTAGCCCTACTCCAGGCCCTTCAAGCCCTGTGAGTACCT 672
DB 1542 GCCACACCCCAAGTGTCCAGTGTAGCCCTACTCCAGGCCCTTCAAGCCCTGTGAGTACCT 1601
QY 673 CTTTGAAGCTGGGGCATCCGCTGGCCGTGTGGGCCATCGTGTGCTCTCCGTGCTCTG 732
DB 1602 CTTTGAAGCTGGGGCATCCGCTGGCCGTGTGGGCCATCGTGTGCTCTCCGTGCTCTG 1661
QY 733 CAATGGACTGTGTGCTGACCGTGTTCGCTGGGGGCTGCCCCCTGCCCCCGGTCAA 792
DB 1662 CAATGGACTGTGTGCTGACCGTGTTCGCTGGGGGCTGCCCCCTGCCCCCGGTCAA 1721
QY 793 GTTGTGTAGTGTGGATGTGAGCGCGCAACACCTTGA CTGGCATTTCTGTGGCTTCT 852
DB 1722 GTTGTGTAGTGTGGATGTGAGCGCGCAACACCTTGA CTGGCATTTCTGTGGCTTCT 1781
QY 853 AGCCTCAGTCAGTGCCTGACCTTTGTGTGCTGCTGAGTACGAGGACCCCTGGAGAC 912
DB 1782 AGCCTCAGTCAGTGCCTGACCTTTGTGTGCTGCTGAGTACGAGGACCCCTGGAGAC 1841
QY 913 GGGGTAGCTGCGGGCCACTGGCTTCTCGCAGTACTTGGCTCGAGGACATCGGTGCT 972
DB 1842 GGGGTAGCTGCGGGCCACTGGCTTCTCGCAGTACTTGGCTCGAGGACATCGGTGCT 1901
QY 973 GCTGCTCACTTGGCGCAGTGCAGTGCAGGGTCTCGGTCTCTGTGTCCGGGCTATGG 1032
DB 1902 GCTGCTCACTTGGCGCAGTGCAGTGCAGGGTCTCGGTCTCTGTGTCCGGGCTATGG 1961
QY 1033 GAAGTCCCCCTTCTGGCAGCGTTCAGCAGGGGTCTAGGCTGCCTGGCAGCTGGCAGG 1092
DB 1962 GAAGTCCCCCTTCTGGCAGCGTTCGAGCAGGGGTCTTAGGCTGCCTGGCAGCTGGCAGG 2021
QY 1093 GCTGGCGCGGCACCTGCCCCCTGGCTCAGTGGGAGATACGGGGCTCCCCACCTCTGCCT 1152
DB 2022 GCTGGCGCGGCACCTGCCCCCTGGCTCAGTGGGAGATACGGGGCTCCCCACCTCTGCCT 2081
QY 1153 GCCCTACGCGCACCTGAGGGTTCAGCCAGCAGCCCTTGGGCTTCAACGCTGGCCCTGGTGA 1212
DB 2082 GCCCTACGCGCACCTGAGGGTTCAGCCAGCAGCCCTTGGGCTTCAACGCTGGCCCTGGTGA 2141
QY 1213 GATGAATCTCTTCTGTTTCTGCTGCTGGCCGCTGCTACATCAAACTGATGTGACCT 1272
DB 2142 GATGAATCTCTTCTGTTTCTGCTGCTGGCCGCTGCTACATCAAACTGATGTGACCT 2201
QY 1273 GCCGCGGGGCACCTTTCAGGGCCGTGGGACTGCGCCATGTGAGGACGCTGGCCCTGGCT 1332
DB 2202 GCCGCGGGGCACCTTTCAGGGCCGTGGGACTGCGCCATGTGAGGACGCTGGCCCTGGCT 2261
QY 1333 CATCTTCGACAGCGGCTCTCTACTGTCCGCTGGCCCTTCTCAGCTTCGCTCCATGCT 1392
DB 2262 CATCTTCGACAGCGGCTCTCTACTGTCCGCTGGCCCTTCTCAGCTTCGCTCCATGCT 2321
QY 1393 GGGCTCTTCTCCTGTACGCCCGAGCGCTCAAGTCTGTCTCTGCTGTGTGCTGCGCCCT 1452
DB 2322 GGGCTCTTCTCCTGTACGCCCGAGCGCTCAAGTCTGTCTCTGCTGTGTGCTGCGCCCT 2381
QY 1453 GCCTGCTGCTCAACCCACTGCTGTACCTCTCTTCAACCCCTCCTTCCGGGATGACCT 1512

DB 2382 GCCTGCTGCTCAACCCACTGCTGTACTGCTTCTCAACCCCTCCTCCGGATGACCT 2441
QY 1513 TCGCGCGCTTTCGGCCCCCGCAGGGGACTCAGGGCCCCCTAGCCTATGCTGCGCGCGGGGA 1572
DB 2442 TCGCGCGCTTTCGGCCCCCGCAGGGGACTCAGGGCCCCCTAGCCTATGCTGCGCGCGGGGA 2501
QY 1573 GCTGGAGAAGAGCTTCCTGATTTCTACCCAGGCCCTGTAGCCTTCTCTGATGTGGATCT 1632
DB 2502 GCTGGAGAAGAGCTTCCTGATTTCTACCCAGGCCCTGTAGCCTTCTCTGATGTGGATCT 2561
QY 1633 CATTTCTGGAAGCTTCTGAAGCTGGCGGCCCTTGGGCTGGAGACCTATGCTTCCCTC 1692
DB 2562 CATTTCTGGAAGCTTCTGAAGCTGGCGGCCCTTGGGCTGGAGACCTATGCTTCCCTC 2621
QY 1693 AGTGAACCTCATCTCTGTGTCAGCAGCCAGGGGCCCCCAGGCTGGAGGCGACCATTTGTGT 1752
DB 2622 AGTGAACCTCATCTCTGTGTCAGCAGCCAGGGGCCCCCAGGCTGGAGGCGACCATTTGTGT 2681
QY 1753 AGAGCCAGAGGGGAACCACTTTTGGGAACCCCAACCTCCATGGATGGAGAACCTGTGCT 1812
DB 2682 AGAGCCAGAGGGGAACCACTTTTGGGAACCCCAACCTCCATGGATGGAGAACCTGTGCT 2741
QY 1813 GAGGCGAGAGGGGATCTACGCCAGCAGGTGGAGGCTTCTCAGGGGCTGGCGCTTTACGCC 1872
DB 2742 GAGGCGAGAGGGATCTACGCCAGCAGGTGGAGGCTTCTCAGGGGCTGGCGCTTTACGCC 2801
QY 1873 CTCTGGCTTGGCTTGTGCTTCAACGCTGTAATATCTCTCCCATTTCTTCTTCCCTC 1932
DB 2802 CTCTGGCTTGGCTTGTGCTTCAACGCTGTAATATCTCTCCCATTTCTTCTTCCCTC 2861
QY 1933 TCTTCCCTTCTCTCTCCCTCGGTGAATGATGCTGCTCTTAAACAAATAACAACCA 1992
DB 2862 TCTTCCCTTCTCTCTCCCTCGGTGAATGATGCTGCTCTTAAACAAATAACAACCA 2921
QY 1993 AAACCTCAGCAGTGTGATCTATAGCAGGATGGCCAGCTACCTGGCTCCACTGATCACCTCT 2052
DB 2922 AAACCTCAGCAGTGTGATCTATAGCAGGATGGCCAGCTACCTGGCTCCACTGATCACCTCT 2981
QY 2053 CTCTGTGACATCACCAACGGGTGCTCTTGGCTGCTTCTTCCCTTGGGCTTCCCTCAGC 2112
DB 2982 CTCTGTGACATCACCAACGGGTGCTCTTGGCTGCTTCTTCCCTTGGGCTTCCCTCAGC 3041
QY 2113 TTCACCTTGTACTGGGCTCTTCTTGTCTATGCTGCTGAAGCTGGGACACAGACCTGGA 2172
DB 3042 TTCACCTTGTACTGGGCTCTTCTTGTCTATGCTGCTGAAGCTGGGACACAGACCTGGA 3101
QY 2173 CTTTGTCTCTTAAAGGAAATGAGGGAAGTAAAGACAGTGAAGGGGTGGAGGTTGATC 2232
DB 3102 CTTTGTCTCTTAAAGGAAATGAGGGAAGTAAAGACAGTGAAGGGGTGGAGGTTGATC 3161
QY 2233 AGGCGACAGTGGACAGGGACCTCAGAGAAAGGCTGGAAGGTGATTTCCCGTGTGA 2292
DB 3162 AGGCGACAGTGGACAGGGACCTCAGAGAAAGGCTTCAAGAGGCTGGAAGGTGATTTCCCGTGTGA 3221
QY 2293 CTATGGATAGGATACAAAATGTTCCTCATGTACCTTAACTTTGACATATGCCATGCAT 2352
DB 3222 CTATGGATAGGATACAAAATGTTCCTCATGTACCTTAACTTTGACATATGCCATGCAT 3281
QY 2353 AAAGACTTCTCTATTAATAAAGCTTTTGGAAAGAGATTAATAAAAAA 2396
DB 3282 AAAGACTTCTCTATTAATAAAGCTTTTGGAAAGAGAAATAAAAAA 3325

RESULT 8

US-10-398-036-26
; Sequence 26, Application US/10398036
; Publication No. US20040137564A1
; GENERAL INFORMATION:
; APPLICANT: INCYTE GENOMICS INC.; BAUGHN, Mariah R.;
; APPLICANT: GRAUL Richard C.; CHAWLA, Narinder K.;
; APPLICANT: GANDHI, Ameena R.; HAFALIA, April J.A.;
; APPLICANT: RAMKUMAR, Jayalaxmi; TRIBOULEY, Catherine M.;


```
Db 2577 CATTCTGGAAGCTTCTGAAGCTGGGCGGCCCTGGGCTGGAGACTATGCTTCCCTC 2636
Qy 1693 AGTGACCTCATCTCTGTACGACGACGGGGCCCCAGGCTGGAGGCGACCATTTGTT 1752
Db 2637 AGTGACCTCATCTCTGTACGACGACGGGGCCCCAGGCTGGAGGCGACCATTTGTT 2696
Qy 1753 AGAGCCAGGGGAACCACTTTGGGAACCCCAACCTCCATGGATGGAGAACTGCTGCT 1812
Db 2697 AGAGCCAGGGGAACCACTTTGGGAACCCCAACCTCCATGGATGGAGAACTGCTGCT 2756
Qy 1813 GAGGCGAGAGGATCTACGCCAGCAGGTGGAGGCTTGTACGGGGGTGGCGCTTTACGCC 1872
Db 2757 GAGGCGAGAGGATCTACGCCAGCAGGTGGAGGCTTGTACGGGGGTGGCGCTTTACGCC 2816
Qy 1873 CTCTGGCTTGGCTTGTCTTACACGCTGAATATCCCTCCCATCTTCTCTCTCCCTC 1932
Db 2817 CTCTGGCTTGGCTTGTCTTACACGCTGAATATCCCTCCCATCTTCTCTCTCCCTC 2876
Qy 1933 TCTTCCCTTCTCTCTCCCTCGGTGAATGATGGCTTCTTAAACAATAACAACA 1992
Db 2877 TCTTCCCTTCTCTCTCCCTCGGTGAATGATGGCTTCTTAAACAATAACAACA 2936
Qy 1993 AAACCTACGAGTGTATCTATAGCAGGATGGCCAGTACCTGGCTCCACTGATCACCTCT 2052
Db 2937 AAACCTACGAGTGTATCTATAGCAGGATGGCCAGTACCTGGCTCCACTGATCACCTCT 2996
Qy 2053 CTCTGTGACCATCACCAAGGGTGCCTTCTGGGCTGGCTTCTTCCCTTGGCTTCTCAGC 2112
Db 2997 CTCTGTGACCATCACCAAGGGTGCCTTCTGGGCTGGCTTCTTCCCTTGGCTTCTCAGC 3056
Qy 2113 TTCACTTGTACTGGGCTCTTCTCTGTCTGAAGTCTGAAGTCTGGACGACGACCTGGA 2172
Db 3057 TTCACTTGTACTGGGCTCTTCTCTGTCTGAAGTCTGGACGACGACCTGGA 3116
Qy 2173 CTTTTGTCTGTTAAGGGAATGAGGGAATGAAAGACAGTGAAGGGGTGGAGGTTGATC 2232
Db 3117 CTTTTGTCTGTTAAGGGAATGAGGGAATGAAAGACAGTGAAGGGGTGGAGGTTGATC 3176
Qy 2233 AGGCGACAGTGGACAGGAGACCTCACAGAAAGGCTGGAAGGTGATTTCCCGTGTGA 2292
Db 3177 AGGCGACAGTGGACAGGAGACCTCACAGAAAGGCTGGAAGGTGATTTCCCGTGTGA 3236
Qy 2293 CTATGATAGATACAAAATGTGTTCCATGTACCATTAATCTTGACATATGCCATGCAT 2352
Db 3237 CTATGATAGATACAAAATGTGTTCCATGTACCATTAATCTTGACATATGCCATGCAT 3296
Qy 2353 AAAGACTTCTTATTAATAAGCTTTGGAAGAGATTAAAA 2393
Db 3297 AAAGACTTCTTATTAATAAGCTTTGGAAGAGATTACACA 3337
```

RESULT 9
us-10-331-496A-94
; Sequence 94, Application US/10331496A
; Publication No. US20030228305A1
; GENERAL INFORMATION:
; APPLICANT: FRANTZ, GRETCHEN
; APPLICANT: HILLAN, KENNETH J.
; APPLICANT: PHILLIPS, HEIDI S.
; APPLICANT: POLAKIS, PAUL
; APPLICANT: SMITH, VICTORIA
; APPLICANT: SPENCER, SUSAN D.
; APPLICANT: WILLIAMS, P. MICKEY
; APPLICANT: WU, THOMAS D.
; APPLICANT: ZHANG, ZEMIN
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE DIAGNOSIS AND
; TREATMENT OF TUMOR
; FILE REFERENCE: P5014P1-PCT
; CURRENT APPLICATION NUMBER: US/10/331,496A
; CURRENT FILING DATE: 2002-12-30
; PRIOR APPLICATION NUMBER: US 60/345,444
; PRIOR FILING DATE: 2002-01-02

```
; PRIOR APPLICATION NUMBER: US 60/351,885  
; PRIOR FILING DATE: 2002-01-25  
; PRIOR APPLICATION NUMBER: US 60/360,066  
; PRIOR FILING DATE: 2002-02-25  
; PRIOR APPLICATION NUMBER: US 60/362,004  
; PRIOR FILING DATE: 2002-03-05  
; PRIOR APPLICATION NUMBER: US 60/366,869  
; PRIOR FILING DATE: 2002-03-20  
; PRIOR APPLICATION NUMBER: US 60/366,284  
; PRIOR FILING DATE: 2002-03-21  
; PRIOR APPLICATION NUMBER: US 60/368,679  
; PRIOR FILING DATE: 2002-03-28  
; PRIOR APPLICATION NUMBER: US 60/404,809  
; PRIOR FILING DATE: 2002-08-19  
; PRIOR APPLICATION NUMBER: US 60/405,645  
; PRIOR FILING DATE: 2002-08-21  
; NUMBER OF SEQ ID NOS: 95  
; SEQ ID NO 94  
; LENGTH: 3443  
; TYPE: DNA  
; ORGANISM: Homo sapien  
US-10-331-496A-94  
  
Query Match 90.0%; Score 2238; DB 17; Length 3443;  
Best Local Similarity 99.6%; Pred. No. 0;  
Matches 2246; Conservative 7; Mismatches 2; Indels 1; Gaps 1;  
  
Qy 133 AGCGGCTCCACACACACCCGATCTGGGAAATTTGGAGCTNGACACCTTCAGCCAGCTGAG 192  
Db 1189 ATCGGCTCCACACACACCCGATCTGGGAAATTTGGAGCT -GACACCTTCAGCCAGCTGAG 1247  
Qy 193 CTCCTCTCAAGCCCTGGATCTTAGCTGGAACGCGCATCCGGTCCATCCACCTCGAGGCTT 252  
Db 1248 CTCCTCTCAAGCCCTGGATCTTAGCTGGAACGCGCATCCGGTCCATCCACCTCGAGGCTT 1307  
Qy 253 CTCACCTCTCACTCCCTGCTCAAGCTGGAACCTGACAGACAAACAGCTGACACCTGCC 312  
Db 1308 CTCACCTCTCACTCCCTGCTCAAGCTGGAACCTGACAGACAAACAGCTGACACCTGCC 1367  
Qy 313 CTGCGCTGGACTTTGGGGCTTGGATCTGAACTCAAGGNAACCTTGTCTCTCTCCA 372  
Db 1368 CTGCGCTGGACTTTGGGGCTTGGATCTGAACTCAAGGNAACCTTGTCTCTCTCCA 1427  
Qy 373 GGCTTCTCCAAAGACAGTTTCCAAAACTGAGGATCTCGAGGTGCTTATGCTTACCA 432  
Db 1428 GGCTTCTCCAAAGACAGTTTCCAAAACTGAGGATCTCGAGGTGCTTATGCTTACCA 1487  
Qy 433 GTGCTGTCCCTATGGGATGTGCCAGCTTCTTCAAGGCTCTTGGGAGTGGGAGGCTGA 492  
Db 1488 GTGCTGTCCCTATGGGATGTGCCAGCTTCTTCAAGGCTCTTGGGAGTGGGAGGCTGA 1547  
Qy 493 AGACCTTCACCTTGATGATGAGGATCTTCAAAAAGGCCCTTGGGCTCTTGGCCAGCA 552  
Db 1548 AGACCTTCACCTTGATGATGAGGATCTTCAAAAAGGCCCTTGGGCTCTTGGCCAGCA 1607  
Qy 553 AGCAGAGAACCACTATGACAGGACCTGGATGAGCTCCAGCTGGAGATGGAGGACTCAA 612  
Db 1608 AGCAGAGAACCACTATGACAGGACCTGGATGAGCTCCAGCTGGAGATGGAGGACTCAA 1667  
Qy 613 GCCACACCCCAAGTGTCCAGTGTAGCCCTACTCCAGGCCCCCTTCAAGCCCTGTGAGTACCT 672  
Db 1668 GCCACACCCCAAGTGTCCAGTGTAGCCCTACTCCAGGCCCCCTTCAAGCCCTGTGAGTACCT 1727  
Qy 673 CTTTGAAGCTGGGGCATCCGCTGGCCGTGTGGGCCATCGTGTGTCTCTCCGTGCTCTG 732  
Db 1728 CTTTGAAGCTGGGGCATCCGCTGGCCGTGTGGGCCATCGTGTGTCTCTCCGTGCTCTG 1787  
Qy 733 CAATGGAAGCTGCTGCTGACCGTGTTCGCTGGCGGCCCTGCCCTCTGCCCGCTCAA 792  
Db 1788 CAATGGAAGCTGCTGCTGACCGTGTTCGCTGGCGGCCCTGCCCTCTGCCCGCTCAA 1847  
Qy 793 GTTGTGTAGGTGCGATTCAGGCGCCAAACA CTTTGACTGGCATTTCTGTGGCTTCT 852  
|||||
```


1848	Db	GTTTGTGGTAGGTGCGAATTGAGCGGCCAACCTTGACTGGCAATTTCTCTGTGGCCCTTCT	1907
853	Qy	AGCCTCAGTCCGATCCCTGACCTTTTGGTCAGTTCTCTGAGTACGGAGCCCGCTGGAGAC	912
1908	Db	AGCCTCAGTCCGATCCCTGACCTTTTGGTCAGTTCTCTGAGTACGGAGCCCGCTGGAGAC	1967
913	Qy	GGGCTTAGGCTGCGCGGCCACTGGCTTCTGGGCAGTACTTTGGTTCGAGGCATCGGTGCT	972
1968	Db	GGGCTTAGGCTGCGCGGCCACTGGCTTCTTGAGTACTTTGGTTCGAGGCATCGGTGCT	2027
973	Qy	GCTGCTCACTCTGSCCGCAGTGCAGCGTCTCCGTCTCTCTGTCGCGGCCCTATGG	1032
2028	Db	GCTGCTCACTCTGSCCGCAGTGCAGCGTCTCCGTCTCTCTGTCGCGGCCCTATGG	2087
1033	Qy	GAAGTCCCCCTCCCTGGGCGACGTTTCAGAGAGGGGTCTTAGGTGCTCGTGCATCTGGCAGG	1092
2088	Db	GAAGTCCCCCTCCCTGGGCGACGTTTCAGAGAGGGGTCTTAGGTGCTCGTGCATCTGGCAGG	2147
1093	Qy	GCTGGCCGCGCACTGCCCCCTGGCTCAGTGGGAGAAATAGCGGCGCTTCCCACTCTGCGCT	1152
2148	Db	GCTGGCCGCGCACTGCCCCCTGGCTCAGTGGGAGAAATAGCGGCGCTTCCCACTCTGCGCT	2207
1153	Qy	GCCCTACGCGCCACTCAGAGGTGACGAGAGCCCTGGGTTCACCGTGGCCCTGGTGAT	1212
2208	Db	GCCCTACGCGCCACTCAGAGGTGACGAGAGCCCTGGGTTCACCGTGGCCCTGGTGAT	2267
1213	Qy	GATGAACCTCTTCTGTTTCTTGTCGTGGCGGTGCTTACATCAAACTGTACTGTGACCT	1272
2268	Db	GATGAACCTCTTCTGTTTCTTGTCGTGGCGGTGCTTACATCAAACTGTACTGTGACCT	2327
1273	Qy	GCGCGGGGCGAATTGAGCCGCTGTCGGGACTGCGCCATGGTAGGCACTGGTGCCTGGCT	1332
2328	Db	GCGCGGGGCGAATTGAGCCGCTGTCGGGACTGCGCCATGGTAGGCACTGGTGCCTGGCT	2387
1333	Qy	CATCTTCGACAGCGGGTCTCTACTGTCGCTGGCCCTTCTCAGCTTCGCTCCATGCT	1392
2388	Db	CATCTTCGACAGCGGGTCTCTACTGTCGCTGGCCCTTCTCAGCTTCGCTCCATGCT	2447
1393	Qy	GGGCTCTTCCCTGTACGCGCCGAGGCCGTCAAGTCTGCTCTGCTGGTGTGTGTCGCCCT	1452
2448	Db	GGGCTCTTCCCTGTACGCGCCGAGGCCGTCAAGTCTGCTCTGCTGGTGTGTGTCGCCCT	2507
1453	Qy	GCTGCTGCTCTCAACCCACTGCTGTACTGCTCTTCAACCCCACTTCGCGGATGACCT	1512
2508	Db	GCTGCTGCTCTCAACCCACTGCTGTACTGCTCTTCAACCCCACTTCGCGGATGACCT	2567
1513	Qy	TCGGCGCTTCGGCCCCGCGCAGGGGACTCAGGGCCCTAGCCTATGCTCGCGCGGGGA	1572
2568	Db	TCGGCGCTTCGGCCCCGCGCAGGGGACTCAGGGCCCTAGCCTATGCTCGCGCGGGGA	2627
1573	Qy	GCTGGAGAAGAGCTCTGTGATTTCTACCCAGGCCCTGCTGAGCTTCTCATGTGATCT	1632
2628	Db	GCTGGAGAAGAGCTCTGTGATTTCTACCCAGGCCCTGCTGAGCTTCTCATGTGATCT	2687
1633	Qy	CATTCTGGAAGCTTCTGAAGCTGGCGGCCCTCTGGCTGGAGACCTATGGCTTCCCTC	1692
2688	Db	CATTCTGGAAGCTTCTGAAGCTGGCGGCCCTCTGGCTGGAGACCTATGGCTTCCCTC	2747
1693	Qy	AGTGACCTCATCTCTGTGACAGCCAGGGGCCCTCAGGGCTGGAGGCACTGTTGT	1752
2748	Db	AGTGACCTCATCTCTGTGACAGCCAGGGGCCCTCAGGGCTGGAGGCACTGTTGT	2807
1753	Qy	AGAGCCAGAGGGGAACCACTTTGGGAACCCCAACCTTCATGGATGGAGAACTGTGCT	1812
2808	Db	AGAGCCAGAGGGGAACCACTTTGGGAACCCCAACCTTCATGGATGGAGAACTGTGCT	2867
1813	Qy	GAGGCGAGAGGATCTACGCCAGCAGGTGGAGGCTTGTCAAGGGGGTGGCGGCTTCAGCC	1872
2868	Db	GAGGCGAGAGGATCTACGCCAGCAGGTGGAGGCTTGTCAAGGGGGTGGCGGCTTCAGCC	2927
1873	Qy	CTCTGGCTTGGCTTTCCTTCAACACGTGTAAATATCCCTCCCACTTCTCTTCCCTC	1932
2928	Db	CTCTGGCTTGGCTTTCCTTCAACACGTGTAAATATCCCTCCCACTTCTCTTCCCTC	2987

Qy	1933	TCCTCCCTTCCTCTCTCCCTCCGCGTGAATGATGGCTGCTCTTAAACAAATACAAACCA	1992
Db	2988	TCCTCCCTTCCTCTCTCCCTCCGCGTGAATGATGGCTGCTCTTAAACAAATACAAACCA	3047
Qy	1993	AAACTCAGCAGTGTGATCTATAGCAGGATGGCCACGATCCTGGCTCCACCTGATCACCCTCT	2052
Db	3048	AAACTCAGCAGTGTGATCTATAGCAGGATGGCCACGATCCTGGCTCCACCTGATCACCCTCT	3107
Qy	2953	CTCCTGTGACATCAACCAACGGGTGCCCTTTGGCTGGCTTTCCCTTTGGCTTCCTCAGC	2112
Db	3108	CTCCTGTGACATCAACCAACGGGTGCCCTTTGGCTGGCTTTCCCTTTGGCTTCCTCAGC	3167
Qy	2113	TTCACTTTGATACTGGGCCCTTCCTTGTCATCTGCTGAAGCTCTGGACACAGACCTTGA	2172
Db	3168	TTCACTTTGATACTGGGCCCTTCCTTTGTCATCTGCTGAAGCTCTGGACACAGACCTTGA	3227
Qy	2173	CTTTTGTCTGCTTAAAGGAAATGAGGAACTAAAGACAGTGAAGGGGTGAGGGTTGATC	2232
Db	3228	CTTTTGTCTGCTTAAAGGAAATGAGGAACTAAAGACAGTGAAGGGGTGAGGGTTGATC	3287
Qy	2233	AGGCACAGTGGACAGGGAGACCTCACAGAGAAAGGCCCTGGAGAGTGATTTCCCGTGTGA	2292
Db	3288	AGGCACAGTGGACAGGGAGACCTCACARAAAAGGCCCTGGAAAGGAGTATTTCCCGTGTGA	3347
Qy	2293	CTCATGTATAGGATACAAAATGTGTTCCATGTACCAATTAATCTTTGACATATGCCATGCAT	2352
Db	3348	CTCATGRTAGGAWACAAAATGTGTTCCATGTACCAATTAATCTTTGACATATGCCATGCAT	3407
Qy	2353	AAAGACTTCCTATTTAAATAAGCTTTTGGAAAGAGATT	2388
Db	3408	AAARACTTCCTATTTAAATAAGCTTTTGGAGAGATT	3443

RESULT 10

```

US-09-851-595-10
; Sequence 10, Application US/09851595
; Publication No. US20030166047A1
; GENERAL INFORMATION:
; APPLICANT: Gu, Wei
; TITLE OF INVENTION: NOVEL G-PROTEIN COUPLED RECEPTORS AND USES THEREFOR
; FILE REFERENCE: WNI-080CP
; CURRENT APPLICATION NUMBER: US/09/851,595
; CURRENT FILING DATE: 2000-05-08
; PRIOR APPLICATION NUMBER: 09/556,588
; PRIOR FILING DATE: 2000-05-08
; PRIOR APPLICATION NUMBER: 60/132,896
; PRIOR FILING DATE: 1999-05-06
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 10
; LENGTH: 3492
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (104)...(3004)
US-09-851-595-10

```

	Query Match	90.0%	Score 2238;	DB 10;	Length 3492;
	Best Local Similarity	99.6%;	Pred. No. 0;		
	Matches 2246;	Conservative	7;	Mismatches	2;
				Indels	1;
				Gaps	1;
Qy	133	AGCGGCTTCAACACAAACCGCATCTGGAAATTGGAGCTNGACACTTCAGCCAGCTGAG	192		
Db	1238	ATCGGCTTCAACACAAACCGCATCTGGAAATTGGAGCT-GACACCTTCAGCCAGCTGAG	1296		
Qy	193	CTCCCTGCAAGCCCTGGATCTTAGCTGGAACGCCATCCGGTCCATCCACCTGAGGCCTT	252		
Db	1297	CTCCCTGCAAGCCCTGGATCTTAGCTGGAACGCCATCCGGTCCATCCACCTGAGGCCTT	1356		
Qy	253	CTCCACCTTCGACTCCCTGGTCAAGCTGGAGCTTGACAGACAACACGAGTGCACCACTGCC	312		

1357	CTCCACCCCTGCACCTCCCTGGTCAAGCTGGACCTTGACAGACAAACAGCTGACCAACACTGGCC	1411
Qy	CCTGGCTGGACTTGGGGGCTTGTATGCATCTGAAGCTCAAAAGGGAACCTTGTCTCTCCCA	372
Db	CCTGGCTGGACTTGGGGGCTTGTATGCATCTGAAGCTCAAAAGGGAACCTTGTCTCTCCCA	1417
Qy	GGCCTTCTTCAAAGACAGTTCCTCCAAACCTGAGGATCTCGGAGGTGCTTATGCTTACCA	432
Db	GGCCTTCTTCAAAGACAGTTCCTCCAAACCTGAGGATCTCGGAGGTGCTTATGCTTACCA	1536
Qy	GTGCTGTCCCTATCGGATGTGTGCGACGCTTCTTCAAGGCCCTCTCGGCAGTGGGAGGCTGA	492
Db	GTGCTGTCCCTATCGGATGTGTGCGACGCTTCTTCAAGGCCCTCTCGGCAGTGGGAGGCTGA	1596
Qy	AGACCTTCACCTTCGATGATGAGGAGTCTTCAAAAAGGCCCTCGGGCTCTCTTGGCAGACA	552
Db	AGACCTTCACCTTCGATGATGAGGAGTCTTCAAAAAGGCCCTCGGGCTCTCTTGGCAGACA	1656
Qy	AGCAGAGAACCACTATGACCAAGGACCTTGGATGAGCTCCAGCTGGAGATGGAGGACTCAAA	612
Db	AGCAGAGAACCACTATGACCAAGGACCTTGGATGAGCTCCAGCTGGAGATGGAGGACTCAAA	1716
Qy	GCCACACCCAGTGTCCAGTGTAGCCCTACTCCAGAGGCCCTTCAAGGCCCTGTGAGTACCT	672
Db	GCCACACCCAGTGTCCAGTGTAGCCCTACTCCAGAGGCCCTTCAAGGCCCTGTGAGTACCT	1776
Qy	CTTTGAAAGCTGGGGCATCCGCCCTGGCGCTGTGGGCCATCGTGTGCTCTCCGTCGCTG	732
Db	CTTTGAAAGCTGGGGCATCCGCCCTGGCGCTGTGGGCCATCGTGTGCTCTCCGTCGCTG	1836
Qy	CAATGGACTGTGTGCTGACCCGTGTTTCGCTGGCGGCCCTGCCCCCTGCCCCCGGTCAA	792
Db	CAATGGACTGTGTGCTGACCCGTGTTTCGCTGGCGGCCCTGCCCCCTGCCCCCGGTCAA	1896
Qy	GTTTGTGGTAGGTGCGAATTGACGGCGCCAAACACTTGACTGGCAATTCCTGTGGCCCTTCT	852
Db	GTTTGTGGTAGGTGCGAATTGACGGCGCCAAACACTTGACTGGCAATTCCTGTGGCCCTTCT	1956
Qy	AGCCTCAGTCGATGCCCTTGACTTTGTGTCAGTTCCTGAGTACGGAGCCCGCTGGGAGAC	912
Db	AGCCTCAGTCGATGCCCTTGACTTTGTGTCAGTTCCTGAGTACGGAGCCCGCTGGGAGAC	2016
Qy	GGGCTTAGGCTCGCGGCCACTGGCTTCTTGCGCAAGTACTTTGGGTTCGGAGGCATCGGTGCT	972
Db	GGGCTTAGGCTCGCGGCCACTGGCTTCTTGCGCAAGTACTTTGGGTTCGGAGGCATCGGTGCT	2076
Qy	GCTGCTCACTCTGGCCGAGTGCAGTGAGGAGTCTCCGTCCTCTGTCCTGGGCGCTATGG	1032
Db	GCTGCTCACTCTGGCCGAGTGCAGTGAGGAGTCTCCGTCCTCTGTCCTGGGCGCTATGG	2136
Qy	GNAGTCCCTCCCTGGGACGTTTCAGCAGGGGTCTTAGGCTGCTGGCACTGGCAGG	1092
Db	GNAGTCCCTCCCTGGGACGTTTCAGCAGGGGTCTTAGGCTGCTGGCACTGGCAGG	2196
Qy	GCTGGCCGCGCACTGCGCCCTGGCTCAGTGGGAGAAATAGGGGCCCTCCCACTCTGCCT	1152
Db	GCTGGCCGCGCACTGCGCCCTGGCTCAGTGGGAGAAATAGGGGCCCTCCCACTCTGCCT	2256
Qy	GCCCTACGCGCCACTGAGGGTACGACAGAGCCCTGGGCTTCAACGTCGCCCTGTGAT	1212
Db	GCCCTACGCGCCACTGAGGGTACGACAGAGCCCTGGGCTTCAACGTCGCCCTGTGAT	2316
Qy	GATGAACTCCTTCTGTTTCTGTTGTCGGGTGCGGTGCCTACATCAAACTGTGTGACCT	1272
Db	GATGAACTCCTTCTGTTTCTGTTGTCGGGTGCGGTGCCTACATCAAACTGTGTGACCT	2376
Qy	GCCGCGGGGCGACTTTTGAGGCCGTGTGGGACTGGGCCATGGTAGGCACTGGGCTGGCT	1332
Db	GCCGCGGGGCGACTTTTGAGGCCGTGTGGGACTGGGCCATGGTAGGCACTGGGCTGGCT	2436
Qy	CATCTTCGACAGCGGCTCTCTACTGTCCGTCGGGCTTCTCCTCAGCTTCGCCCTCCATGCT	1392
Db	CATCTTCGACAGCGGCTCTCTACTGTCCGTCGGGCTTCTCCTCAGCTTCGCCCTCCATGCT	2496

Qy	1393	GGGCGCTTCTCCCTGTGTACAGCCCGAGGCGGTCAAGTCTGTCTCTGCTGGTGGTGTGCTGCCCT	1455
Db	2497	GGGCGCTTCTCCCTGTGTACAGCCCGAGGCGGTCAAGTCTGTCTCTGCTGGTGGTGTGCTGCCCT	2556
Qy	1453	GCTTGCCTGCTCAAACCCACTGCTGTACCTGCTCTTCAACCCCACTTCCGGAGTAGACCT	1512
Db	2557	GCTTGCCTGCTCAAACCCACTGCTGTACCTGCTCTTCAACCCCACTTCCGGAGTAGACCT	2616
Qy	1513	TCGGCGGCTTTCGGCCCCGGCGCAGGGGACTCAGGGCCCCCTAGCCTATATGCTCGCGCGGGGA	1572
Db	2617	TCGGCGGCTTTCGGCCCCGGCGCAGGGGACTCAGGGCCCCCTAGCCTATATGCTCGCGCGGGGA	2676
Qy	1573	GCTGGAGAAAGACTCTCTGTATTTCTACCAAGGCCCTGGTAGCCTTCTCTGATGTGGATCT	1632
Db	2677	GCTGGAGAAAGACTCTCTGTATTTCTACCAAGGCCCTGGTAGCCTTCTCTGATGTGGATCT	2736
Qy	1633	CATTCTGGAAGCTTCTGAAGCTGGCGGGCCCCCTGGGCTGGAGACCTATGGCTTCCCCCTC	1692
Db	2737	CATTCTGGAAGCTTCTGAAGCTGGCGGGCCCCCTGGGCTGGAGACCTATGGCTTCCCCCTC	2796
Qy	1693	AGTGACCCCTCATCTCTGTGTAGCAGCCAGGGGCCCCCAGGCTGGAGGGCAGCAATTTGTGT	1752
Db	2797	AGTGACCCCTCATCTCTGTGTAGCAGCCAGGGGCCCCCAGGCTGGAGGGCAGCAATTTGTGT	2856
Qy	1753	AGAGCCAGAGGGAAACCACTTTTGGGAACCCCAACCCCTCCATGTGATGGAGAACTGCTGCT	1812
Db	2857	AGAGCCAGAGGGAAACCACTTTTGGGAACCCCAACCCCTCCATGTGATGGAGAACTGCTGCT	2916
Qy	1813	GAGGGCAGAGGGATCTACGCCAGCAGGTGGAGGCTTTGTTCAGGGGGTGGCGGCTTTTCAGCC	1872
Db	2917	GAGGGCAGAGGGATCTACGCCAGCAGGTGGAGGCTTTGTTCAGGGGGTGGCGGCTTTTCAGCC	2976
Qy	1873	CTCTGGCTTGGCCCTTGTCTTCAACGTGTAAATATATCCCTCCCACTTCTTCTCTTCCCTC	1932
Db	2977	CTCTGGCTTGGCCCTTGTCTTCAACGTGTAAATATATCCCTCCCACTTCTTCTCTTCCCTC	3036
Qy	1933	TCCTTCCCTTTCCTCTCTCCCCCTCGGTGAATGATGCTGCTTCTTAAACAAATACAAACA	1992
Db	3037	TCCTTCCCTTTCCTCTCTCCCCCTCGGTGAATGATGCTGCTTCTTAAACAAATACAAACA	3096
Qy	1993	AAACTCAGCAGTGTGATCTATAGCAGGATGGCCAGTACTCGCTCCACTGATCACTCT	2052
Db	3097	AAACTCAGCAGTGTGATCTATAGCAGGATGGCCAGTACTCGCTCCACTGATCACTCT	3156
Qy	2053	CTCCTGTGACCATCAACACGGGTGCTCTTGGCCTGGGCTTTCCCTTGGGCTTCTCTCAGC	2112
Db	3157	CTCCTGTGACCATCAACACGGGTGCTCTTGGCCTGGGCTTTCCCTTGGGCTTCTCTCAGC	3216
Qy	2113	TTCACTTTTGATCTAGGGGCTTCTCTTGTTCATGCTCTGAAGCTGTGGACACAGAGACTGGA	2172
Db	3217	TTCACTTTTGATCTAGGGGCTTCTCTTGTTCATGCTCTGAAGCTGTGGACACAGAGACTGGA	3276
Qy	2173	CTTTTGTCTGTAAAGGAAATGAGGGAAGTAAAGA CAGTGAAGGGGTGGAGGGTGTGATC	2232
Db	3277	CTTTTGTCTGTAAAGGAAATGAGGGAAGTAAAGA CAGTGAAGGGGTGGAGGGTGTGATC	3336
Qy	2233	AGGGCAGGTGGACAGGGGAGCCTCACAGAGAAAGCCCTGGAAGGTGATTTCCCGTGTGA	2292
Db	3337	AGGGCAGGTGGACAGGGGAGCCTCACAGAGAAAGCCCTGGAAGGTGATTTCCCGTGTGA	3396
Qy	2293	CTCATGGATAGGATACAAATGTGTTTCCATGTACCAATTAATCTTTGACATATGCCATGCAT	2352
Db	3397	CTCATGGRTAGGAWACAAATGTGTTTCCATGTACCAATTAATCTTTGACATATGCCATGCAT	3456
Qy	2353	AAAGACTTCTCTATTTAAATAAGCTTTTGGAAAGATTT	2388
Db	3457	AAAPACTTCTCTATTTAAATAAGCTTTTGGAGAGATTT	3492

```

; Publication No. US20040059377A1
; GENERAL INFORMATION:
; APPLICANT: Gu, Wei
; TITLE OF INVENTION: NOVEL G-PROTEIN COUPLED RECEPTORS AND USES THEREFOR
; FILE REFERENCE: WNI-080CP
; CURRENT APPLICATION NUMBER: US/10/664,667
; CURRENT FILING DATE: 2003-09-18
; PRIOR APPLICATION NUMBER: US/09/851,595
; PRIOR FILING DATE: 2000-05-08
; PRIOR APPLICATION NUMBER: 09/556,588
; PRIOR FILING DATE: 2000-05-08
; PRIOR APPLICATION NUMBER: 60/132,896
; PRIOR FILING DATE: 1999-05-06
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 10
; LENGTH: 3492
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (104)...(3004)
US-10-664-667-10

Query Match          90.0%; Score 2238; DB 17; Length 3492;
Best Local Similarity 99.6%; Pred. No. 0;
Matches 2246; Conservative 7; Mismatches 2; Indels 1; Gaps 1;

Qy      133  AGCGGCTTCAACACCAACCCGATCTGGGAAATTGGAGCTNGACACCTTCAGCCAGCTGAG 192
Db      1238  ATCGGCTTCAACACCAACCCGATCTGGGAAATTGGAGCT-GACACCTTCAGCCAGCTGAG 1596

Qy      193  CTCCTTGCAAGCCCTGGATCTTAGCTGGAACGCCATCCGGTCCATCCACCTCAGGCGCTT 252
Db      1297  CTCCTTGAAGCCCTGGATCTTAGCTGGAACGCCATCCGGTCCATCCACCTGAGGCGCTT 1356

Qy      253  CTCACCCCTGGACTCCCTGGTCAAGCTGGACCTGACAGACAAACAGGTGACCAACATGACC 312
Db      1357  CTCACCCCTGGACTCCCTGGTCAAGCTGGACCTGACAGACAAACAGGTGACCAACATGACC 1416

Qy      313  CTTGGCTGGACTTGGGGGCTTGATGATCTGAGCTCAAGCTCAAAAGGNAACCTTCTCTCCCA 372
Db      1417  CTTGGCTGGACTTGGGGGCTTGATGATCTGAGCTCAAGCTCAAAAGGNAACCTTCTCTCCCA 1476

Qy      373  GGCCTTCTCAAAGACAGATTTCCTCAAACCTGAGGATCCTGGAGTGCTTTATGCTACCA 432
Db      1477  GGCCTTCTCAAAGACAGATTTCCTCAAACCTGAGGATCCTGGAGTGCTTTATGCTACCA 1536

Qy      433  GTGCTGTCTTATGGGATGTGTGCCAGCTTCTTCAAGCCTCTGTGGCAGTGGGAGGCTGA 492
Db      1537  GTGCTGTCTTATGGGATGTGTGCCAGCTTCTTCAAGCCTCTGTGGCAGTGGGAGGCTGA 1596

Qy      493  AGACCTTCACCTTGATGATGAGGAGTCTTCAAAAAGCCCCCTGGGCTCTTGTCCAGACA 552
Db      1597  AGACCTTCACCTTGATGATGAGGAGTCTTCAAAAAGCCCCCTGGGCTCTTGTCCAGACA 1656

Qy      553  AGCAGAAACCACTATGACCAAGACCTGGATGAGTCCAGCTGAGATGAGAGACTCAAA 612
Db      1657  AGCAGAAACCACTATGACCAAGACCTGGATGAGTCCAGCTGAGATGAGAGACTCAAA 1716

Qy      613  GCCACACCCAGTGTCCAGTGTAGCCCTACTCCAGGCCCTTCAAGCCCTGTGAGTACCT 672
Db      1717  GCCACACCCAGTGTCCAGTGTAGCCCTACTCCAGGCCCTTCAAGCCCTGTGAGTACCT 1776

Qy      673  CTTTGAAGCTGGGGCATTCGCCCTGGCCGTGTGGGCCCATCGTGTGTCTTCCGTGCTGTG 732
Db      1777  CTTTGAAGCTGGGGCATTCGCCCTGGCCGTGTGGGCCCATCGTGTGTCTTCCGTGCTGTG 1836

Qy      733  CAATGGACTGGTGTCTGTACCGGTGTTGCTGGCGGCCCTGCCCCCTGCCCGGTCAA 792
Db      1837  CAATGGACTGGTGTCTGTACCGGTGTTGCTGGCGGCCCTGCCCCCTGCCCGGTCAA 1896

Qy      793  GTTTGTGCTGAGTGCAGATTGACGCGCCAAACACCTTGACTGGCAATTTCTGTGGCGCTTCT 852

```

Db 2977 CTCTGGCTGGCTTTGGCTTACACAGTGTAATATATCCCTCCCATCTTCTCTCTCCCTC 3036
 Qy 1933 TCTTCCCTTTCTCTCTCCCTCGGTGAATGAGTGGCTGCTTAAACAAATACACCA 1992
 Db 3037 TCTTCCCTTTCTCTCTCCCTCGGTGAATGAGTGGCTGCTTAAACAAATACACCA 3096
 Qy 1993 AAATCTCAGCAGTGTGATCTATAGCAGGATGCCAGTACCTGGCTCCACTGATCACCTCT 2052
 Db 3097 AAATCTCAGCAGTGTGATCTATAGCAGGATGCCAGTACCTGGCTCCACTGATCACCTCT 3156
 Qy 2053 CTCTGTGACCATCACCAACGGGTCTCTTGTCTATGCTGAGCTGTGGACGAGACCTTGA 2172
 Db 3157 CTCTGTGACCATCACCAACGGGTCTCTTGTCTATGCTGAGCTGTGGACGAGACCTTGA 3216
 Qy 2113 TTCACTTGTACTTGGGCTCTTCTTGTCTATGCTGAGCTGTGGACGAGACCTTGA 2172
 Db 3217 TTCACTTGTACTTGGGCTCTTCTTGTCTATGCTGAGCTGTGGACGAGACCTTGA 3276
 Qy 2173 CTTTGTCTGCTTAAGGGAATGAGGAAGTAAAGACAGTGAAGGGTGGAGGCTTGATC 2232
 Db 3277 CTTTGTCTGCTTAAGGGAATGAGGAAGTAAAGACAGTGAAGGGTGGAGGCTTGATC 3336
 Qy 2233 AGGCACAGTGGACAGGAGACCTCACAGAGAAAGGCTGGAAGGTGATTTCCCGTGTGA 2292
 Db 3337 AGGCACAGTGGACAGGAGACCTCACAGAGAAAGGCTGGAAGGTGATTTCCCGTGTGA 3396
 Qy 2293 CTCTAGTATAGGATACAAATGTGTTCCATGATACCATTAATCTTGACATATGCCATGAT 2352
 Db 3397 CTCTAGTATAGGATACAAATGTGTTCCATGATACCATTAATCTTGACATATGCCATGAT 3456
 Qy 2353 AAAGACTTCTTATTAATAAATAGCTTTGGAAGAGATT 2388
 Db 3457 AAARACTTCTTATTAATAAAGCTTTGGRAGATT 3492

RESULT 12

US-10-737-450-31
 ; Sequence 31, Application US/10737450
 ; Publication No. US20040235071A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Millennium Pharmaceuticals, Inc
 ; APPLICANT: Lightcap, Eric S.
 ; APPLICANT: Ecsedy, Jeffrey A.
 ; APPLICANT: Hunter, John Joseph
 ; APPLICANT: MacBeth, Kyle J.
 ; APPLICANT: Tighe Nestor, Michelle
 ; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR TREATING
 ; TITLE OF INVENTION: CANCER USING 15986, 2186, 20743, 9148, 9151, 9791, 44252,
 ; TITLE OF INVENTION: 14184, 42461, 8204, 7970, 25552, 21657, 26492, 2411, 15088,
 ; TITLE OF INVENTION: 1905, 28899, 63380, 33935, 10480, 12686, 25501, 17694,
 ; TITLE OF INVENTION: 15701, 53062, 49908, 21612, 38949, 6216, 46863, 9235, 2201,
 ; TITLE OF INVENTION: 6985, 9883, 12238, 18057, 21617, 39228, 49928, 54476, 62113,
 ; TITLE OF INVENTION: 64316, 12264, 32362, 58198, 2887, 3205, 8557, 9600, 9693,
 ; TITLE OF INVENTION: 44867, 53058, 55556, 57658, 2208, 10252, 10302, 14218,
 ; TITLE OF INVENTION: 33877, 10317, 10485, 25964, 14815, 1363, 1397, 14827, 21708,
 ; TITLE OF INVENTION: 3801, 64698, 2179 OR 13249
 ; FILE REFERENCE: MPI02-207PIRNONIM
 ; CURRENT APPLICATION NUMBER: US/10737,450
 ; CURRENT FILING DATE: 2003-12-16
 ; PRIOR FILING DATE: 2003-12-16
 ; PRIOR FILING DATE: 2002-12-20
 ; PRIOR FILING DATE: 2002-12-20
 ; PRIOR FILING DATE: 2002-12-23
 ; PRIOR FILING DATE: 2003-01-07
 ; PRIOR FILING DATE: 2003-01-31
 ; PRIOR FILING DATE: 2003-01-31
 ; PRIOR FILING DATE: 2003-02-06
 ; PRIOR FILING DATE: 2003-03-11
 ; PRIOR FILING DATE: 2003-03-11
 ; PRIOR FILING DATE: 2003-03-25

; PRIOR APPLICATION NUMBER: US 60/462,458
 ; PRIOR FILING DATE: 2003-04-10
 ; PRIOR APPLICATION NUMBER: US 60/466,732
 ; PRIOR FILING DATE: 2003-04-30
 ; PRIOR APPLICATION NUMBER: US 60/469,184
 ; PRIOR FILING DATE: 2003-05-08
 ; Remaining Prior Application data removed - See File Wrapper or PALM.
 ; NUMBER OF SEQ ID NOS: 144
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 31
 ; LENGTH: 3492
 ; TYPE: DNA
 ; ORGANISM: Homo Sapiens
 ; FEATURE:
 ; NAME/KEY: CDS
 ; LOCATION: (104)...(3007)
 US-10-737-450-31
 Query Match 90.0%; Score 2238; DB 18; Length 3492;
 Best Local Similarity 99.6%; Pred. No. 0;
 Matches 2246; Conservative 7; Mismatches 2; Indels 1; Gaps 1;
 Qy 133 AGCGGCTCCAAACACCAACCCATCTGGGAAATTTGAGCTNGACACCTTCAGCCAGCTGAG 192
 Db 1238 ATCGGCTCCAAACACCAACCCATCTGGGAAATTTGAGCT-GACACCTTCAGCCAGCTGAG 1296
 Qy 193 CTGCTGCAAGCCCTGGATCTTAGCTGGAAGCCATCCGGTCCATCCACCTGAGGCTTT 252
 Db 1297 CTGCTGCAAGCCCTGGATCTTAGCTGGAAGCCATCCGGTCCATCCACCTGAGGCTTT 1356
 Qy 253 CTCACCGCTCACCTCCCTGGTCAAGCTGGAGCTGACAGACCAACAGCTGACCAACACCTGCC 312
 Db 1357 CTCACCGCTCACCTCCCTGGTCAAGCTGGAGCTGACAGACCAACAGCTGACCAACACCTGCC 1416
 Qy 313 CTGCTGCAAGCCCTGGATCTTAGCTGGAAGCCATCCGGTCCATCCACCTGAGGCTTT 372
 Db 1417 CTGCTGCAAGCCCTGGATCTTAGCTGGAAGCCATCCGGTCCATCCACCTGAGGCTTT 1476
 Qy 373 GCGCTTCTCAAGGACAGTTTCCCAAACTGAGATCTCTGGAGTGGCTTATGCTTACCA 432
 Db 1477 GCGCTTCTCAAGGACAGTTTCCCAAACTGAGATCTCTGGAGTGGCTTATGCTTACCA 1536
 Qy 433 GTGCTGCTCCATATGGGATGTGCTCAGCTTCTTCAAGGCTCTTGGGAGTGGAGGCTGA 492
 Db 1537 GTGCTGCTCCATATGGGATGTGCTCAGCTTCTTCAAGGCTCTTGGGAGTGGAGGCTGA 1596
 Qy 493 AGACCTTTCACCTTGTATGATGAGGAGTCTTCAAAAAGGCCCTCTGGGCTCTCTTGCACACA 552
 Db 1597 AGACCTTTCACCTTGTATGATGAGGAGTCTTCAAAAAGGCCCTCTGGGCTCTCTTGCACACA 1656
 Qy 553 AGCAGAGAACCTATGACAGGACCTGGATGAGCTCCAGCTGGAGATGGAGGACTCAAA 612
 Db 1657 AGCAGAGAACCTATGACAGGACCTGGATGAGCTCCAGCTGGAGATGGAGGACTCAAA 1716
 Qy 613 GCCACACCCAGTGTCCAGTGTAGCCCTACTTCCAGGCCCCCTTCAAGGCCCTGTGAGTACCT 672
 Db 1717 GCCACACCCAGTGTCCAGTGTAGCCCTACTTCCAGGCCCCCTTCAAGGCCCTGTGAGTACCT 1776
 Qy 673 CTTTGAAGCTGGGGCATCCGCTGGCCGTGTGGGCCATCGTGTGTCTCTCGGTCTGT 732
 Db 1777 CTTTGAAGCTGGGGCATCCGCTGGCCGTGTGGGCCATCGTGTGTCTCTCGGTCTGT 1836
 Qy 733 CAATGACCTGGTCTGCTGACCGTGTGCTGGGGGCCCTGCCCTCCCTCCCGGCTCAA 792
 Db 1837 CAATGACCTGGTCTGCTGACCGTGTGCTGGGGGCCCTGCCCTCCCTCCCGGCTCAA 1896
 Qy 793 GTTTGTGAGTGGTCCGATTTGACGCGCAACACCTTGACCTGGCATTTCTGTGCGCTTCT 852
 Db 1897 GTTTGTGAGTGGTCCGATTTGACGCGCAACACCTTGACCTGGCATTTCTGTGCGCTTCT 1956
 Qy 853 AGCCTCAGTCGATGCCCTGACCTTTGGTTCAGTGTCTCTGAGTACGAGCCCGCTGGAGAC 912
 Db 1957 AGCCTCAGTCGATGCCCTGACCTTTGGTTCAGTGTCTCTGAGTACGAGCCCGCTGGAGAC 2016

QY 913 GGGCTAGGCTCCGGGCACTGGCTTCTGCGAGTACTTGGGTGGAGGCATCGGTGCT 972
Db 2017 GGGCTAGGCTCCGGGCACTGGCTTCTGCGAGTACTTGGGTGGAGGCATCGGTGCT 2076
QY 973 GGTGCTCACTCTGGCGGAGTGCAGTGCAGTCTCCGCTCTCTGTCGGGCTATGG 1032
Db 2077 GGTGCTCACTCTGGCGGAGTGCAGTGCAGTCTCCGCTCTCTGTCGGGCTATGG 2136
QY 1033 GAAGTCCCGCTCCCTGGGCACTGGCTTCTGCGAGGCTTCTGAGGCTGCTGGCACTGGCAGG 1092
Db 2137 GAAGTCCCGCTCCCTGGGCACTGGCTTCTGCGAGGCTTCTGAGGCTGCTGGCACTGGCAGG 2196
QY 1093 GGTGCGGCGCACTGCTCCCTGGCTCACTGAGGAGAAATACGGGGCTTCCCACTGTGCT 1152
Db 2197 GGTGCGGCGCACTGCTCCCTGGCTCACTGAGGAGAAATACGGGGCTTCCCACTGTGCT 2256
QY 1153 GCGCTACCGGCACTGAGGCTCAGCGACGAGCCTGGGCTTCACGGTGGGCTGGTAT 1212
Db 2257 GCGCTACCGGCACTGAGGCTCAGCGACGAGCCTGGGCTTCACGGTGGGCTGGTAT 2316
QY 1213 GATGAATCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1272
Db 2317 GATGAATCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2376
QY 1273 GCGCGGCGCACTTGGAGCGCTGTGGCACTGCGCCATGTGGAGCACTGGCTGCTGCT 1332
Db 2377 GCGCGGCGCACTTGGAGCGCTGTGGCACTGCGCCATGTGGAGCACTGGCTGCTGCT 2436
QY 1333 CATCTTCGAGAGGGCTCTCTACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1392
Db 2437 CATCTTCGAGAGGGCTCTCTACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2496
QY 1393 GGGCTCTTCCCTGTCACGCCGAGCGCTCAAGTCTGCTGCTGCTGCTGCTGCTGCT 1452
Db 2497 GGGCTCTTCCCTGTCACGCCGAGCGCTCAAGTCTGCTGCTGCTGCTGCTGCTGCT 2556
QY 1453 GCTGCTGCTGCTCAACCCCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1512
Db 2557 GCTGCTGCTGCTCAACCCCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2616
QY 1513 TCGCGGCTTCCGCGCGGCAAGGCACTCAGGCGGCTTCTGCTGCTGCTGCTGCTGCT 1572
Db 2617 TCGCGGCTTCCGCGCGGCAAGGCACTCAGGCGGCTTCTGCTGCTGCTGCTGCTGCT 2676
QY 1573 GCTGAGAGAGCTCTGCTGATTTCTACCCAGGCTGCTGCTGCTGCTGCTGCTGCT 1632
Db 2677 GCTGAGAGAGCTCTGCTGATTTCTACCCAGGCTGCTGCTGCTGCTGCTGCTGCT 2736
QY 1633 CATCTGGAAGCTTCTGAAGCTGGGCGGCTTCTGCTGCTGCTGCTGCTGCTGCTGCT 1692
Db 2737 CATCTGGAAGCTTCTGAAGCTGGGCGGCTTCTGCTGCTGCTGCTGCTGCTGCTGCT 2796
QY 1693 AGTGACCTCATCTCTGTCAGCAGCAGGCGGCTTCTGCTGCTGCTGCTGCTGCTGCT 1752
Db 2797 AGTGACCTCATCTCTGTCAGCAGCAGGCGGCTTCTGCTGCTGCTGCTGCTGCTGCT 2856
QY 1753 AGAGCAGAGGGGAACCACTTTGGGAACCCCAACCTCCATGATGGAGAACTGCTGCT 1812
Db 2857 AGAGCAGAGGGGAACCACTTTGGGAACCCCAACCTCCATGATGGAGAACTGCTGCT 2916
QY 1813 GAGGCGAGAGGATCTACGCCAGCAGGTGGAGGCTTGTGAGGGGTGGGCTTTCAGCC 1872
Db 2917 GAGGCGAGAGGATCTACGCCAGCAGGTGGAGGCTTGTGAGGGGTGGGCTTTCAGCC 2976
QY 1873 CTCTGGCTTGGCTTCTGCTTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1932
Db 2977 CTCTGGCTTGGCTTCTGCTTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3036
QY 1933 TCTTCCCTTCTCTCTCTCCCTCGGTGATGCTGCTTCTTAAACAAATACACCA 1992
Db 3037 TCTTCCCTTCTCTCTCTCCCTCGGTGATGCTGCTTCTTAAACAAATACACCA 3096

RESULT 13

US-10-295-027-1119
; Sequence 1119, Application US/10295027
; Publication No. US20030232350A1
; GENERAL INFORMATION:
; APPLICANT: Afar, Daniel
; APPLICANT: Aziz, Natasha
; APPLICANT: Ginsberg, Wendy M.
; APPLICANT: Gish, Kurt C.
; APPLICANT: Glynn, Richard
; APPLICANT: Hevezi, Peter A.
; APPLICANT: Mack, David H.
; APPLICANT: Murray, Richard
; APPLICANT: Watson, Susan R.
; APPLICANT: Eos Biotechnology, Inc.
; TITLE OF INVENTION: Methods of Diagnosis of Cancer, Compositions and
; FILE REFERENCE: 018501-012500US
; CURRENT APPLICATION NUMBER: US/10/295,027
; CURRENT FILING DATE: 2002-11-13
; PRIOR APPLICATION NUMBER: US 09/663,733
; PRIOR FILING DATE: 2000-09-15
; PRIOR APPLICATION NUMBER: US 60/350,666
; PRIOR FILING DATE: 2001-11-13
; PRIOR APPLICATION NUMBER: US 60/335,394
; PRIOR FILING DATE: 2001-11-15
; PRIOR APPLICATION NUMBER: US 60/332,464
; PRIOR FILING DATE: 2001-11-21
; PRIOR APPLICATION NUMBER: US 60/334,393
; PRIOR FILING DATE: 2001-11-29
; PRIOR APPLICATION NUMBER: US 60/340,376
; PRIOR FILING DATE: 2001-12-14
; PRIOR APPLICATION NUMBER: US 60/347,211
; PRIOR FILING DATE: 2002-01-08
; PRIOR APPLICATION NUMBER: US 60/347,349
; PRIOR FILING DATE: 2002-01-10
; PRIOR APPLICATION NUMBER: US 60/355,250
; PRIOR FILING DATE: 2002-02-08
; PRIOR APPLICATION NUMBER: US 60/356,714
; PRIOR FILING DATE: 2002-02-13
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 1386
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1119

; LENGTH: 3273									
; TYPE: DNA									
; ORGANISM: Homo sapiens									
US-10-295-027-1119									
Query Match 89.8%; Score 2231.4; DB 17; Length 3273;									
Best Local Similarity 99.7%; Pred. No. 0;									
Matches 2246; Conservative 0; Mismatches 6; Indels 1; Gaps 1;									
Qy	133	AGCGCCCTCCAAACACACCGCATCTGGGAATTTGGAGCTNGACACTTTCAGGCGAGCTGAG	192						
Db	1022	ATCGGCTCCCAACACACACCGCATCTGGGAATTTGGAGCT-GACACCTTTCAGGCGAGCTGAG	1080						
Qy	193	CTCCCTGCAAGCCCTGGATCTTAGCTGGAAAGCCCATCGGTCCATCCACCCCGAGGCCCTT	252						
Db	1081	CTCCCTGCAAGCCCTGGATCTTAGCTGGAAAGCCCATCGGTCCATCCACCCCGAGGCCCTT	1140						
Qy	253	CTCCACCTCTGACCTCCCTGGTCAAGCTGGACCTGACAGACAACCAAGCTGACCACTGCC	312						
Db	1141	CTCCACCTCTGACCTCCCTGGTCAAGCTGGACCTGACAGACAACCAAGCTGACCACTGCC	1200						
Qy	313	CCTGCTGGACCTTGGGGCTTGATGCACTCTGAAGCTCAAGGAAACCTTGTCTCTCTCCCA	372						
Db	1201	CCTGCTGGACCTTGGGGCTTGATGCACTCTGAAGCTCAAGGAAACCTTGTCTCTCTCCCA	1260						
Qy	373	GGCTTCTCCAAAGGACAGTTTCCAAAACCTGAGGATCTCTGGAGGTGCCCTTATGCTTACCA	432						
Db	1261	GGCTTCTCTCCAAAGGACAGTTTCCAAAACCTGAGGATCTCTGGAGGTGCCCTTATGCTTACCA	1320						
Qy	433	GTGCTGTCCCTATGGGATGTGTGCAGCTTCTTCAAGGCCCTCTGGGCAGTGGGAGGCTGA	492						
Db	1321	GTGCTGTCCCTATGGGATGTGTGCAGCTTCTTCAAGGCCCTCTGGGCAGTGGGAGGCTGA	1380						
Qy	493	AGACCTTCACTTGTATGATGAGGAGTCTTCAAAAAGGCCCTTGGGCCCTCTCTTGCACACA	552						
Db	1381	AGACCTTCACTTGTATGATGAGGAGTCTTCAAAAAGGCCCTTGGGCCCTCTCTTGCACACA	1440						
Qy	553	AGCAGAGAACCACTATGACACAGGACCTCGATGAGCTCCAGCTGGAGATGGAGGACTCAAA	612						
Db	1441	AGCAGAGAACCACTATGACACAGGACCTCGATGAGCTCCAGCTGGAGATGGAGGACTCAAA	1500						
Qy	613	GCCACACCCCAAGTGTCCAGTGTAGCCCTACTCCAGGCCCTTCAAGGCCCTCTGAGTACT	672						
Db	1501	GCCACACCCCAAGTGTCCAGTGTAGCCCTACTCCAGGCCCTTCAAGGCCCTCTGAGTACT	1560						
Qy	673	CTTTGAAAGCTGGGGCATCCGCTGGCGTGTGGGCCATCGTGTGTCTCCGTGCTGTG	732						
Db	1561	CTTTGAAAGCTGGGGCATCCGCTGGCGTGTGGGCCATCGTGTGTCTCCGTGCTGTG	1620						
Qy	733	CAATGGAGTGGTGTGCTGACCGTCTTGGCTGGCGGCCCTGCCCTGCCCGGTCAA	792						
Db	1621	CAATGGAGTGGTGTGCTGACCGTCTTGGCTGGCGGCCCTGTGCCCTGCCCGGTCAA	1680						
Qy	793	GTTTGTGTAGTGTGATGTCAGGCGGCCAAACACTTGACTGGCATTTCTGTGGCCCTTCT	852						
Db	1681	GTTTGTGTAGTGTGATGTCAGGCGGCCAAACACTTGACTGGCATTTCTGTGGCCCTTCT	1740						
Qy	853	AGCTCTAGTGCATGCCCTGACCTTTGGTCAAGTCTCTGAGTACGGAGCCCGCTGGGAGAC	912						
Db	1741	AGCTCTAGTGCATGCCCTGACCTTTGGTCAAGTCTCTGAGTACGGAGCCCGCTGGGAGAC	1800						
Qy	913	GGGGCTAGGCTGCCGGGCCACTGGCTTCTGGCAGTACTTGGGTCCGAGGCACTCGGTCT	972						
Db	1801	GGGGCTAGGCTGCCGGGCCACTGGCTTCTGGCAGTACTTGGGTCCGAGGCACTCGGTCT	1860						
Qy	973	GCTGCTCACTGTGGCGCAGTGCAGGTCTCCGCTCTCTGTGTGTCTGGGGCTATGG	1032						
Db	1861	GCTGCTCACTGTGGCGCAGTGCAGGTCTCCGCTCTCTGTGTGTCTGGGGCTATGG	1920						
Qy	1033	GAAGTCCCTCTCCCTGGGACGGTTCGAGCAGGGGTCTTAGGCTCCCTGGCACTGGCAGG	1092						
Db	1921	GAAGTCCCTCTCCCTGGGACGGTTCGAGCAGGGGTCTTAGGCTCCCTGGCACTGGCAGG	1980						
Qy	1093	GCTGGCGCGCACTGCCCCTGGGCTCAGTGGGAGAAATACGGGGCCTCCCCACACTCTGCCT	1152						
Db	1981	GCTGGCGCGCGCTGCCCCTGGGCTCAGTGGGAGAAATACGGGGCCTCCCCACACTCTGCCT	2040						
Qy	1153	GCCCTACGCGCCACCTGAGGGTCAGCAGCAGCCCTGGGCTTCAACGTTGGCCCTTGGTAT	1212						
Db	2041	GCCCTACGCGCCACCTGAGGGTCAGCAGCAGCCCTGGGCTTCAACGTTGGCCCTTGGTAT	2100						
Qy	1213	GATGAACTCCTTCTGTTTCTGTTGCTGGCGGTGCTTACATCAAACTGATCTGACCT	1272						
Db	2101	GATGAACTCCTTCTGTTTCTGTTGCTGGCGGTGCTTACATCAAACTGATCTGACCT	2160						
Qy	1273	GCCCGGGGACCTTTCAGGCGCTGTGGGACTGTGGCCATGTGTAGAGGACGCTGGGCTGGCT	1332						
Db	2161	GCCCGGGGACCTTTCAGGCGCTGTGGGACTGTGGCCATGTGTAGAGGACGCTGGGCTGGCT	2220						
Qy	1333	CATCTTCGCGAGAAGGGCTCTCTACTGTCCCGTGGCCTTCTCTCAGCTTTCGCTCCATGCT	1392						
Db	2221	CATCTTCGCGAGAAGGGCTCTCTACTGTCCCGTGGCCTTCTCTCAGCTTTCGCTCCATGCT	2280						
Qy	1393	GGGCTCTTTCCTGTACGCGCCGAGCGGCTCAAGTCTCTCTGCTGGTGGTGTGCCCT	1452						
Db	2281	GGGCTCTTTCCTGTACGCGCCGAGCGGCTCAAGTCTCTCTGCTGGTGGTGTGCCCT	2340						
Qy	1453	GCCTGCTGCTCAACCACTGCTGTACTGCTCTTCAACCCCACTTCCGGGATGACCT	1512						
Db	2341	GCCTGCTGCTCAACCACTGCTGTACTGCTCTTCAACCCCACTTCCGGGATGACCT	2400						
Qy	1513	TCGGCGCTTTCGGCCCGCGCAGGGGACTCAGGGCCCTTACGCTTATGCTCGGCGCGGGA	1572						
Db	2401	TCGGCGCTTTCGGCCCGCGCAGGGGACTCAGGGCCCTTACGCTTATGCTCGGCGCGGGA	2460						
Qy	1573	GCTGGAGAGAGCTCTCTGATTTACCCAGGCCCTTGTAGCTTCTCTGATGGAATCT	1632						
Db	2461	GCTGGAGAGAGCTCTCTGATTTACCCAGGCCCTTGTAGCTTCTCTGATGGAATCT	2520						
Qy	1633	CATTCTCGAAAGCTTCTGAAGCTGGCGGCCCTTGGGCTGGAGACTTATGGCTTCCCCTC	1692						
Db	2521	CATTCTCGAAAGCTTCTGAAGCTGGCGGCCCTTGGGCTGGAGACTTATGGCTTCCCCTC	2580						
Qy	1693	AGTGACCTCTCATCTCTGTGTCAGCAGCCAGGGGCCCCCAGGCTGGAGGCGACCAATGTGT	1752						
Db	2581	AGTGACCTCTCATCTCTGTGTCAGCAGCCAGGGGCCCCCAGGCTGGAGGCGACCAATGTGT	2640						
Qy	1753	AGAGCCAGAGGGGAAACCACTTTGGGAAACCCCAACCTTCCATGATGGAGACTGCTGCT	1812						
Db	2641	AGAGCCAGAGGGGAAACCACTTTGGGAAACCCCAACCTTCCATGATGGAGACTGCTGCT	2700						
Qy	1813	GAGGGCAGAGGGATCTACGCGCAGCGAGGTGGAGGCTTGTCAAGGGGTGGCGCTTTCAGCC	1872						
Db	2701	GAGGGCAGAGGGATCTACGCGCAGCGAGGTGGAGGCTTGTCAAGGGGTGGCGCTTTCAGCC	2760						
Qy	1873	CTCTGGCTTGGCCCTTTGCTTTCACAGTGAATAATCCCTCCCAATTTCTTCTTCCCCTC	1932						
Db	2761	CTCTGGCTTGGCCCTTTGCTTTCACAGTGAATAATCCCTCCCAATTTCTTCTTCCCCTC	2820						
Qy	1933	TCCTTCCCTTCTCTCTCCCTCTGGTGAATGATGGCTGCTTCTTAAACAAATACACCA	1992						
Db	2821	TCCTTCCCTTCTCTCTCCCTCTGGTGAATGATGGCTGCTTCTTAAACAAATACACCA	2880						
Qy	1993	AACTCAGCAGTGTGATCTATAGCAGGATGCCCGCAGTACCTGGCTCCACTGATCACCTCT	2052						
Db	2881	AACTCAGCAGTGTGATCTATAGCAGGATGCCCGCAGTACCTGGCTCCACTGATCACCTCT	2940						
Qy	2053	CTCCTGTGACCATCACCAACGGGTGCTCTTGGCTGGCTTTTCCCTTGGCTTCCCTCAGC	2112						
Db	2941	CTCCTGTGACCATCACCAACGGGTGCTCTTGGCTGGCTTTTCCCTTGGCTTCCCTCAGC	3000						
Qy	2113	TTCACTTGTACTTGGGCTCTTCTTGTGTATGTCTGAAAGCTGTGGACCAAGAGACTTGA	2172						
Db	3001	TTCACTTGTACTTGGGCTCTTCTTGTGTATGTCTGAAAGCTGTGGACCAAGAGACTTGA	3060						
Qy	2173	CTTTTGTCTCTTAAAGGGAATGAGGGAAGTAAAGACAGTGAAGGGGTGGAGGTTGATC	2232						

Db	2581		AGTGACCTCATCTCTGTGACGACGAGGGCCCCAGAGCTGGAGGCGACGCAATGTGT	2640
Qy	1753	AGAGCCAGAGGGGAACCACTTTGGGAAACCCCAACCCCTCCATGGATGAGAACTGCTGCT	1812	
Db	2641	AGAGCCAGAGGGGAACCACTTTGGGAAACCCCAACCCCTCCATGGATGAGAACTGCTGCT	2700	
Qy	1813	GAGGCGAGAGGATCTAGCCGACGAGGTGGAGGCTTGTTCAGGGGTGGGGCTTTCAGCC	1872	
Db	2701	GAGGCGAGAGGATCTAGCCGACGAGGTGGAGGCTTGTTCAGGGGTGGGGCTTTCAGCC	2760	
Qy	1873	CTCTGGCTTGGCTTTGCTTTCACAGTGTAAATATPCCCTCCCAATTCCTTCTTCCCTC	1932	
Db	2761	CTCTGGCTTGGCTTTGCTTTCACAGTGTAAATATPCCCTCCCAATTCCTTCTTCCCTC	2820	
Qy	1933	TCTTCCCTTCTCTCTCCCTCGGTGAATGATGGCTGTCTTAAACAATAACAACA	1992	
Db	2821	TCTTCCCTTCTCTCTCCCTCGGTGAATGATGGCTGTCTTAAACAATAACAACA	2880	
Qy	1993	AAACTCAGCAGTGTATATAGCAGGATGCCAGTACCTGGCTCCACTGATCACCTCT	2052	
Db	2881	AAACTCAGCAGTGTATATAGCAGGATGCCAGTACCTGGCTCCACTGATCACCTCT	2940	
Qy	2053	CTCTGTGACCATCACCAACGGGTGCTCTTGGGCTGGCTTTCCCTTGGCCCTTCTCAGC	2112	
Db	2941	CTCTGTGACCATCACCAACGGGTGCTCTTGGGCTGGCTTTCCCTTGGCCCTTCTCAGC	3000	
Qy	2113	TTCACTTGTATCTGGGCTCTTCTTGTCTATGTCTGAAGCTGTGGACGAGACCTGGA	2172	
Db	3001	TTCACTTGTATCTGGGCTCTTCTTGTCTATGTCTGAAGCTGTGGACGAGACCTGGA	3060	
Qy	2173	CTTTGTCTGCTTAAGGGAATGAGGAGTAAAGACAGTGAAGGGTGGAGGCTTGATC	2232	
Db	3061	CTTTGTCTGCTTAAGGGAATGAGGAGTAAAGACAGTGAAGGGTGGAGGCTTGATC	3120	
Qy	2233	AGGCGACAGTGGACAGGGAGACCTCACAGAGAAAGCCCTGGAAGTGATTTCCCGTGTGA	2292	
Db	3121	AGGCGACAGTGGACAGGGAGACCTCACAGAGAAAGCCCTGGAAGTGATTTCCCGTGTGA	3180	
Qy	2293	CTCATGATAGGATACAAAATGTGTTCCATGTACATTAATTTTGAATATGCCATGCAT	2352	
Db	3181	CTCATGATAGGATACAAAATGTGTTCCATGTACATTAATTTTGAATATGCCATGCAT	3240	
Qy	2353	AAAGACTTCCTATTAAATATAGCTTTGGAAG	2385	
Db	3241	AAAGACTTCCTATTAAATATAGCTTTGGAAG	3273	
RESULT 15				
US-09-970-944-5				
; Sequence 5, Application US/09970944				
; Publication No. US20030204052A1				
; GENERAL INFORMATION:				
; APPLICANT: Hertman, John L				
; APPLICANT: Rastelli, Luca				
; APPLICANT: Shinkets, Richard A				
; TITLE OF INVENTION: No. US20030204052A1el Proteins and Nucleic Acids Encoding Same an				
; TITLE OF INVENTION: Antibodies Directed Against these Proteins				
; FILE REFERENCE: 21402-138				
; CURRENT APPLICATION NUMBER: US/09/970,944				
; PRIOR FILING DATE: 2002-05-02				
; PRIOR APPLICATION NUMBER: 60/237,862				
; PRIOR FILING DATE: 2000-10-04				
; NUMBER OF SEQ ID NOS: 62				
; SOFTWARE: PatentIn Ver. 2.1				
; SEQ ID NO 5				
; LENGTH: 3381				
; TYPE: DNA				
; ORGANISM: Homo sapiens				
US-09-970-944-5				
Query Match 89.3%; Score 2219.6; DB 10; Length 3381;				
Best Local Similarity 99.7%; Pred. No. 0;				

Matches 2255; Conservative 0; Mismatches 4; Indels 3; Gaps 3;				
Qy	133	AGGGCCCTCCAAACAACCCGATCTGGGAAATTGGAGCTNGACACCTTCACCCAGCTGAG	192	
Db	1058	ATCGGCTTCCAAACAACCCGATCTGGGAAATTGGAGCT-GACACCTTCAGCCAGCTGAG	1116	
Qy	193	CTCCCTCAAGCCCTGATCTTAGCTGAAACGCGCATCCGGTCCATCCACCCCTGAGGCTT	252	
Db	1117	CTCCCTCAAGCCCTGATCTTAGCTGAAACGCGCATCCGGTCCATCCACCCGAGGCTT	1176	
Qy	253	CTCCACCTCGACCTCCCTGGTCAAGCTGGAGCTTGACAGACAACACAGCTGACACACTGCC	312	
Db	1177	CTCCACCTCGACCTCCCTGGTCAAGCTGGAGCTTGACAGACAACACAGCTGACACACTGCC	1236	
Qy	313	CTGGCTGGACTTGGGGCTTGATGATCTGAAGCTCAAAGGAACTTGTCTCTCCCA	372	
Db	1237	CTGGCTGGACTTGGGGCTTGATGATCTGAAGCTCAAAGGAACTTGTCTCTCCCA	1296	
Qy	373	GGCTTCTCCAAGGACAGTTTCCCAAACTGAGGATCTTGAGGTGCTTATGCTTACCA	432	
Db	1297	GGCTTCTCCAAGGACAGTTTCCCAAACTGAGGATCTTGAGGTGCTTATGCTTACCA	1356	
Qy	433	GTGCTGCTCCCTATGGGATGTGTCAGCTTCTTCAAGGCCCTCTGGGAGTGGGAGGTGA	492	
Db	1357	GTGCTGCTCCCTATGGGATGTGTCAGCTTCTTCAAGGCCCTCTGGGAGTGGGAGGTGA	1416	
Qy	493	AGACTTCACTTGAATGATGAGGAGTCTTCAAAGGCCCTTGGGCTCTCTTGGCAGACA	552	
Db	1417	AGACTTCACTTGAATGATGAGGAGTCTTCAAAGGCCCTTGGGCTCTCTTGGCAGACA	1476	
Qy	553	AGCAGAGAAACCACTATGACAGGACCTGGATGAGTCCAGCTGAGATGGAGGACTCAAA	612	
Db	1477	AGCAGAGAAACCACTATGACAGGACCTGGATGAGTCCAGCTGAGATGGAGGACTCAAA	1536	
Qy	613	GCCACACCCAGTGTCCAGTGTAGCCCTACTCCAGGCCCTTCAAGGCCCTGTGAGTACCT	672	
Db	1537	GCCACACCCAGTGTCCAGTGTAGCCCTACTCCAGGCCCTTCAAGGCCCTGTGAGTACCT	1596	
Qy	673	CTTTGAAAGCTGGGCGATCCGCTGGCCGTGTGGGCCATCGTGTGCTCTCCGTGCTGTG	732	
Db	1597	CTTTGAAAGCTGGGCGATCCGCTGGCCGTGTGGGCCATCGTGTGCTCTCCGTGCTGTG	1656	
Qy	733	CAATGGACTCGTCTGCTGACCGTGTGCTGGGGCCCTGCCCCCTGCCCGGTCAA	792	
Db	1657	CAATGGACTCGTCTGCTGACCGTGTGCTGGGGCCCTGCCCCCTGCCCGGTCAA	1716	
Qy	793	GTTTGTGTTAGTGCGATTGACAGCGCCAAACACTTGACTGGCATTTTCTGTGGCTTCT	852	
Db	1717	GTTTGTGTTAGTGCGATTGACAGCGCCAAACACTTGACTGGCATTTTCTGTGGCTTCT	1776	
Qy	853	AGCCTCAGTCGATGCCCTGACCTTTGGTCAAGTCTCTGAGTACGGAGCCCGCTGGGAGAC	912	
Db	1777	AGCCTCAGTCGATGCCCTGACCTTTGGTCAAGTCTCTGAGTACGGAGCCCGCTGGGAGAC	1836	
Qy	913	GGGCTAGGCTGCGGGCCACTGCTTCTCGGAGTACTTGGGTGGAGGCACTCGGTGCT	972	
Db	1837	GGGCTAGGCTGCGGGCCACTGCTTCTCGGAGTACTTGGGTGGAGGCACTCGGTGCT	1896	
Qy	973	GCTGCTCACTCTGCGCGCAGTGCAGTGCAGCGTCTCCGTCTCTGTGTCCGGGCTATGG	1032	
Db	1897	GCTGCTCACTCTGCGCGCAGTGCAGTGCAGCGTCTCCGTCTCTGTGTCCGGGCTATGG	1956	
Qy	1033	GAAGTCCCTCTCCCTGGGCGAGCTTGGAGCAGGGGTCTTAGGCTCGCTGGCAGTGGCAGG	1092	
Db	1957	GAAGTCCCTCTCCCTGGGCGAGCTTGGAGCAGGGGTCTTAGGCTCGCTGGCAGTGGCAGG	2016	
Qy	1093	GCTGGCGCGCAGTGCCTCGCTCAGTGGAGGAAATACGGGGCCCTCCCACTCTCCT	1152	
Db	2017	GCTGGCGCGCAGTGCCTCGCTCAGTGGAGGAAATACGGGGCCCTCCCACTCTCCT	2076	
Qy	1153	GCCTTAGCGCCACCTGAGGGTCAGCCAGCAGCCCTGGGCTTCAACCGTGGCCCTGGTGT	1212	
Db	2077	GCCTTAGCGCCACCTGAGGGTCAGCCAGCAGCCCTGGGCTTCAACCGTGGCCCTGGTGT	2136	

QY 1213 GATGAACCTCCTTCTCTGTTCTTGCTGCTGCGCGGTGCTTACATCAAACTGTACTGTGACCT 1272
Db 2137 GATGAACCTCCTTCTCTGTTCTTGCTGCTGCGCGGTGCTTACATCAAACTGTACTGTGACCT 2196
QY 1273 GCCCGGGGCGACTTTGAGGCGCGTGTGGACGTGCCCATGTGTGAGGCACGTGGCCCTGGCT 1332
Db 2197 GCCCGGGGCGACTTTGAGGCGCGTGTGGACGTGCCCATGTGTGAGGCACGTGGCCCTGGCT 2256
QY 1333 CATCTTCGACAGCGGCTCCTCTACTGTCCCGTGGCTTCCCTCAAGTCTGTCTGCTGGTGGTCTGCCCT 1392
Db 2257 CATCTTCGACAGCGGCTCCTCTACTGTCCCGTGGCTTCCCTCAAGTCTGTCTGCTGGTGGTCTGCCCT 2316
QY 1393 GGGCCTCTTCCCTGTCAAGCCCGAGCGCGCTCAAGTCTGTCTGCTGGTGGTCTGCCCT 1452
Db 2317 GGGCCTCTTCCCTGTCAAGCCCGAGCGCGCTCAAGTCTGTCTGCTGGTGGTCTGCCCT 2376
QY 1453 GCCTGCCCTCAACCCCACTGCTGTCTGCTTCAACCCCACTTCCGGGATGACCT 1512
Db 2377 GCCTGCCCTCAACCCCACTGCTGTCTGCTTCAACCCCACTTCCGGGATGACCT 2436
QY 1513 TCGGCGGCTTCGGCCCGCGAGGGGACTCAGGGCCCTAGGCTATGCTGCGGCCGGGA 1572
Db 2437 TCGGCGGCTTCGGCCCGCGAGGGGACTCAGGGCCCTAGGCTATGCTGCGGCCGGGA 2496
QY 1573 GCTGGAGAAGAGCTCCTGTGATTTACCCAGGCCCTGGTAGCCTTCTCTGATGTGGATCT 1632
Db 2497 GCTGGAGAAGAGCTCCTGTGATTTACCCAGGCCCTGGTAGCCTTCTCTGATGTGGATCT 2556
QY 1633 CATCTGGAAGCTTCTGAAGCTGGCGGCCCTGGCTGGAGACCTATGGCTTCCCTC 1692
Db 2557 CATCTGGAAGCTTCTGAAGCTGGCGGCCCTGGCTGGAGACCTATGGCTTCCCTC 2616
QY 1693 AGTGACCTCATCTCTGTGACAGCAGGGGCCCGCCAGGCTGGAGGCGAGCATTTGT 1752
Db 2617 AGTGACCTCATCTCTGTGACAGCAGGGGCCCGCCAGGCTGGAGGCGAGCATTTGT 2676
QY 1753 AGAGCCAGAGGGGAACCACTTTGGGAACCCCAACCCCTCCATGGATGGAGAACTGCTGCT 1812
Db 2677 AGAGCCAGAGGGGAACCACTTTGGGAACCCCAACCCCTCCATGGATGGAGAACTGCTGCT 2736
QY 1813 GAGGCGAGAGGATCTAGCCAGCAGGTGGAGGCTTGTCAAGGGGTGGCGGCTTCAGCC 1872
Db 2737 GAGGCGAGAGGATCTAGCCAGCAGGTGGAGGCTTGTCAAGGGGTGGCGGCTTCAGCC 2795
QY 1873 CTCTGGCTTGGCTTTGCTTTCACAGTGTAAATATCCCTCCCATTTCTTCTTCCCTC 1932
Db 2796 CTCTGGCTTGGCTTTGCTTTCACAGTGTAAATATCCCTCCCATTTCTTCTTCCCTC 2855
QY 1933 TCTTCCCTTCTCTCTCCCTCGGTGAATGATGGCTGCTTTCTAAACAAATACAACA 1992
Db 2856 TCTTCCCTTCTCTCTCCCTCGGTGAATGATGGCTGCTTTCTAAACAAATACAACA 2915
QY 1993 AAATCAGAGTGTGATCTATAGCAGGATGGCCAGTACCTGGCTCCAATGATCACTCT 2052
Db 2916 AAATCAGAGTGTGATCTATAGCAGGATGGCCAGTACCTGGCTCCAATGATCACTCT 2975
QY 2053 CTCTGTGACCATCAACAGGGTG-CTCTTGGCTGGCTTCCCTTGGCTTCCCTCAG 2111
Db 2976 CTCTGTGACCATCAACAGGGTGCCCTCTTGGCTGGCTTCCCTTGGCTTCCCTCAG 3035
QY 2112 CTTCACTTGTATCTGGGCTCTTCTCTGTCTGATGTCTGAAGCTGTGGACCAAGAGACCTGG 2171
Db 3036 CTTCACTTGTATCTGGGCTCTTCTCTGTCTGATGTCTGAAGCTGTGGACCAAGAGACCTGG 3095
QY 2172 ACTTTTGTCTGTTAAGGGAATGAGGGAAGTAAAGACAGTGAAGGGGTGGAGGTTGAT 2231
Db 3096 ACTTTTGTCTGTTAAGGGAATGAGGGAAGTAAAGACAGTGAAGGGGTGGAGGTTGAT 3155
QY 2232 CAGGCGACAGTGGACAGGGAGACCTCAGAGAAAGGCTGGAAGGTGATTTCCCGTGTG 2291
Db 3156 CAGGCGACAGTGGACAGGGAGACCTCAGAGAAAGGCTGGAAGGTGATTTCCCGTGTG 3215

QY 2292 ACTCATGTAGGATACAAAATGTGTCCATGTACCATTAATCTTTGACATATGCCATGCA 2351
Db 3216 ACTCATGTAGGATACAAAATGTGTCCATGTACCATTAATCTTTGACATATGCCATGCA 3275
QY 2352 TAAAGACTTCCCTATTAAATAAGCTTTGGAAGAGATTAAAA 2393
Db 3276 TAAAGACTTCCCTATTAAATAAGCTTTGGAAGAGATTACACA 3317

Search completed: April 14, 2005, 19:13:35
Job time : 950.376 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: April 14, 2005, 15:59:23 ; Search time 5014 Seconds
(without alignments)
6117.302 Million cell updates/sec

Title: US-10-664-667-5

Perfect score: 3307

Sequence: 1 NTHYRESWACYRSGIPG.....GGLSGGGGFGPSGLAFASHV 633

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 4708233 seqs, 24227607955 residues

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlh
-Q=/cgn2_1/USPTO_epool/US10664667/runat_13042005_184539_7264/app_query_fasta_1.775
-DB=GenEmbl -QPMF=fastap -SUFFIX=p2n.rge -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-LOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000
-USR=US10664667@cgn 1 1 2799 @runat_13042005_184539_7264 -NCPU=6 -ICPU=3
-NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

GenEmbl:*
1: gb_ba:*
2: gb_hcg:*
3: gb_in:*
4: gb_on:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_ats:*
12: gb_sy:*
13: gb_un:*
14: gb_vi:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB ID	Description
1	3305	99.9	1899	6	AX301827 Sequence
2	3305	99.9	2486	6	AX301825 Sequence
3	3150	95.3	2322	6	AX098217 Sequence
4	3140	95.0	3451	6	CQ840814 Sequence

5	3045.5	92.1	2208	6	AX301830
6	3045.5	92.1	2711	6	AX301828
7	3045.5	92.1	2901	6	AX301833
8	3045.5	92.1	3119	9	AX549295
9	3045.5	92.1	3119	9	AB049405
10	3045.5	92.1	3325	9	AX358119
11	3045.5	92.1	3429	6	AX451929
12	3045.5	92.1	3492	6	AX301831
13	3041.5	92.0	3273	6	CQ870634
14	3041.5	92.0	3286	9	BC047905
15	3041.5	92.0	3305	6	CQ840811
16	3040.5	91.9	2487	6	AX926297
17	3038.5	91.9	3349	6	AX926300
18	3016.5	91.2	3041	6	E42917
19	3016.5	91.2	3041	6	AX016191
20	3016.5	91.2	3381	6	AX527920
21	3015.5	91.2	2484	9	AF190501
22	2638	79.8	2901	6	AX301824
23	2638	79.8	3637	6	AX301822
24	2380	72.0	1987	6	E42916
25	2380	72.0	1987	6	AX016189
26	2197	66.4	177344	9	AL356953
27	2193	66.3	119596	6	AX926299
28	2184.5	66.1	1272	9	AB083616
29	2087	63.1	2044	10	BC026896
30	2045	61.8	1681	6	BD159457
31	2045	61.8	1681	6	AX882015
32	2045	61.8	1681	6	BD012484
33	2045	61.8	1681	9	AK027377
34	2030	61.4	178757	2	AC074007
35	1907.5	57.7	253599	2	AC117827
C	1905	57.6	227401	2	AC105613
C	1905	57.6	238582	2	AC098661
37	1905	57.6	247017	2	AC110384
38	1905	57.6	2724	6	CQ721366
39	1275	38.6	2724	6	CQ766844
40	1275	38.6	2724	6	AX308745
41	1275	38.6	2724	6	AX549136
42	1275	38.6	2724	9	AF061444
43	1275	38.6	2724	9	AF061444
44	1275	38.6	2880	6	AX658241
45	1275	38.6	2880	9	AF062006

ALIGNMENTS

RESULT 1	AX301827	AX301827	Sequence 6 from Patent WO0185768.	1899 bp	DNA	linear	PAT 30-NOV-2001
LOCUS	AX301827	AX301827	Sequence 6 from Patent WO0185768.				
DEFINITION	AX301827	AX301827	Sequence 6 from Patent WO0185768.				
ACCESSION	AX301827	AX301827	Sequence 6 from Patent WO0185768.				
VERSION	AX301827.1	AX301827.1	GI:17382885				
KEYWORDS	AX301827.1	AX301827.1	GI:17382885				
SOURCE	AX301827.1	AX301827.1	GI:17382885				
ORGANISM	AX301827.1	AX301827.1	GI:17382885				
REFERENCE	AX301827.1	AX301827.1	GI:17382885				
AUTHORS	AX301827.1	AX301827.1	GI:17382885				
TITLE	AX301827.1	AX301827.1	GI:17382885				
JOURNAL	AX301827.1	AX301827.1	GI:17382885				
FEATURES	AX301827.1	AX301827.1	GI:17382885				
source	AX301827.1	AX301827.1	GI:17382885				
CDS	AX301827.1	AX301827.1	GI:17382885				

Novel g-protein coupled receptors and uses therefor
Patent: WO 0185768-A 6 15-NOV-2001;
Millennium Pharmaceuticals, Inc. (US)
Location/Qualifiers
1. 1899
/organism="Homo sapiens"
/db_xref="taxon:9606"
<1..>1899
/note="unnamed protein product"
/codon_start=1
/protein_id="CAD13078.1"
/db_xref="GI:17382886"
/translation="NTHYRESWACYRSGIPGSTHVSQGLSLFAPHPASLAAL
AASNTTASGKLELDTFSQLSLQALDLSWNAIRSIHPFEAFSTLSLVKLDLTDQLTT

LPLAGLGLMLHLKGNLALSOAFPSKDSFPKLRILEVPYAYQCPYGMCAFFKASQO
 WEARDLHLDDESESKRPLGLILARQAEHNHYDLDLDELQLEMEDSKPHSPVOCSPTEGPF
 KPCEYLPESEGIRLANVAI VLLSVLCNGLVLLTVFAGPGAPLPVPKVFVGAAGANTIL
 TGISCGLIASVDALTFQGFSEYGARWETGLUCRATGFLAVLGSEASVILLTFLAAVQCS
 VVSCVRAYGKSPSLGSRVAGVLGCLALAGLAAALPLASVGEYCASPLCLIFAPPEGO
 PAALGFTVALVMNNSFCFLVVAGAYIKLYCDLPRGDFEAVWDCAMVRHVAWLI PADGL
 LYCPVAFSLFASMLKLPVTPPEAVKSVLLVPLPACLNPLLYLLFNPHRDRRLRR
 PRAGDSPLAVAAAGELEKSCDSTOALVAFSDVDLILLEASEAGRPPGLETYGFPFSVT
 LLSQQQAPAPLESHCVPEPBNHFGNPQPSMDGELLILRAEGSTPAAGSLSGGGGFPQ
 SGLAFASHV"

ORIGIN

Alignment Scores:

Pred. No.: 1,64e-226 Length: 1899
 Score: 3305.00 Matches: 632
 Percent Similarity: 99.84% Conservative: 0
 Best Local Similarity: 99.84% Mismatches: 1
 Query Match: 99.94% Indels: 0
 DB: 6 Gaps: 0

US-10-664-667-5 (1-633) x AX301827 (1-1899)

Qy 1 AenThrThrHisTyrArgGluSerTrpTyrAlaCysArgTyrArSerGlyIleProGly 20
 Db 1 AATACGACTACTATAGGAAAGCTGGTACGCCCTGCAGGTACCGGTCCGGAAATCCCGGG 60
 Qy 21 SerThrHisAlaSerValGluArgSerGlnGlyLeuSerLeuProAlaHisProAlaSer 40
 Db 61 TCGACCCACGCGTCCGTGGAGCGAGCGAGGGTCTGAGCCCTGCGGCTCATCCAGCCCTCT 120
 Qy 41 LeuAlaAlaLeuAlaAlaSerAsnThrThrAlaSerGlyIleLeuGlu***AspThrPhe 60
 Db 121 CTTCGTCGCCCTAGCGGCTCCACACACACGCCATCTGGGAAATTTGGAGCTNGACACCTTC 180
 Qy 61 SerGlnLeuSerSerLeuGlnAlaLeuAspLeuSerTrpAsnAlaIleArgSerIleHis 80
 Db 181 AGCCAGCTGAGCTCCCTGCAAGCCCTGGATCTTAGCTGGAAACGCCATCCGGTCCATCCAC 240
 Qy 81 ProGluAlaPheSerThrLeuHisSerLeuValLysLeuAspLeuThrAspAsnGlnLeu 100
 Db 241 CCTGAGGCTTCTCCACCCCTGCACCTCCCTTGGTCAAGCTGGAGCTGCAGACCAACAGCTG 300
 Qy 101 ThrThrLeuProLeuAlaGlyLeuGlyLeuMetHisLeuLysLeuLysGlyAsnLeu 120
 Db 301 ACCACATGCCCTGGCTGGACTTTGGGGCTTGATGATCTGAAGCTCAAAAGGAACTT 360
 Qy 121 AlaLeuSerGlnAlaPheSerLysAspSerPheProLysLeuArgIleLeuGluValPro 140
 Db 361 GCTCTCTCCAGGCTTCTCCAAAGACAGATTTCCCAAACTGAGGATCTCTGGAGGTGCCT 420
 Qy 141 TyrAlaTyrGlnCysCysProTyrGlyMetCysAlaSerPheLysAlaSerGlyGln 160
 Db 421 TATGCCCTACCACTGCTGCTCCATATGGATGTGTGCCAGCTTCTTCAAGGCTCTTGGGCAG 480
 Qy 161 TrpGluAlaGluAspLeuHisLeuAspGluGluSerSerLysArgProLeuGlyLeu 180
 Db 481 TGGAGGCTGAAGACCTTACCTTGATGATGAGGAGTCTTCAAAAGGCCCTCGGCCTC 540
 Qy 181 LeuAlaArgGlnAlaGluAsnHisTyrAspGlnAspLeuAspGluLeuGlnLeuMet 200
 Db 541 CTTGCCAGACAAAGCAGAGAACCACTATGACCAGGACCTGGATGAGCTCCAGCTGGAGATG 600
 Qy 201 GluAspSerLysProHisProSerValGlnCysSerProThrProGlyProPheLysPro 220
 Db 601 GAGGACTCAAAAGCCACACCCAGGTGCTCAGGTAGCCCTACTCCAGGCCCTTCAAGCCC 660
 Qy 221 CysGlnTyrLeuPheGluSerTrpGlyIleArgLeuAlaValTrpAlaIleValLeuLeu 240
 Db 661 TGTGAGTACCTCTTTGAAGCTGGGGCATCCGCTGGCCGCTGGGCCATCGTGTGCTC 720
 Qy 241 SerValLeuCysAsnGlyLeuValLeuLeuThrValPheAlaGlyIleProAlaProLeu 260
 Db 721 TCCGTGCTCTGCAATGACCTGGTCTGCTGACCGTGTTCGCTGGCGGCTGCCCCCTG 780

Qy 261 ProProValIysPheValValGlyValAlaIleAlaGlyAlaAsnThrLeuThrGlyIleSer 280
 Db 781 CCCCCGGTCAAGTTTGTGTAGGTGGCATTCGACGGCCCAACACTTGTACTGGCATTTCC 840
 Qy 281 CysGlyLeuLeuAlaSerValAspAlaLeuThrPheGlyGlnPheSerGluTyrGlyAla 300
 Db 841 TGTGGCCTTCTAGCCTCAGTCGATGCCCTGACCTTTGTGTCAGTTCTCTGAGTACGGAGCC 900
 Qy 301 ArgTrpGluThrGlyLeuGlyCysArgAlaThrGlyPheLeuAlaValLeuGlySerGlu 320
 Db 901 CGCTGGAGACGGGGCTAGGCTGCCGGGCCACTGGCTTCTTGGCAGTACTTGGGTGCGAG 960
 Qy 321 AlaSerValLeuLeuLeuThrLeuAlaAlaValGlnCysSerValSerValSerCysVal 340
 Db 961 GCATCGGTGCTGCTGCTCACTCTGGCCGCGAGTCAGTGCAGCGTCTCCGTCTCTGTGTC 1020
 Qy 341 ArgAlaTyrGlyIysSerProSerLeuGlySerValArgAlaGlyValLeuGlyCysLeu 360
 Db 1021 CGGGCCTATGGGAAGTCCCTCCCTGGGCGAGGTTCCGACAGGGGTCTTAGGCTGCCGTG 1080
 Qy 361 AlaLeuAlaGlyLeuAlaAlaLeuProLeuAlaSerValGlyGluTyrGlyAlaSer 380
 Db 1081 GCATCGCAGGGCTGGCCGCGCACTGCCCTTGGCCTCAGTGGGAGAATAACGGGGCCTCC 1140
 Qy 381 ProLeuCysLeuProTyrAlaProProGluGlyGlnProAlaLeuGlyPheThrVal 400
 Db 1141 CCACCTCTGCTGCTTACCGCGCACCTGAGGGTTCAGCAGAGGCTTGGGGCTTCCACGTG 1200
 Qy 401 AlaLeuValMetMetAsnSerPheCysPheLeuValValAlaGlyAlaTyrIleLysLeu 420
 Db 1201 GCCCTGGTATGATGAATCTCTTCTTCTTGGTCTGGCCGGTGCCTACATCAAACTG 1260
 Qy 421 TyrCysAspLeuProArgGlyAspPheGluAlaValTrpAspCysAlaMetValArgHis 440
 Db 1261 TACTGTGACCTGCGCGGGCGCATTTTGGCCGCTGTGGGACTGCGGCATGGTGAGGCAC 1320
 Qy 441 ValAlaTrpLeuIlePheAlaAspGlyLeuLeuTyrCysProValAlaPheLeuSerPhe 460
 Db 1321 GTGGCCTGGCTCATCTTTCGACAGCGGGCTCTCTTACTGTCCCCTGGGGCTTCTCTCAGCTTC 1380
 Qy 461 AlaSerMetLeuGlyLeuPheProValThrProGluAlaValLysSerValLeuLeuVal 480
 Db 1381 GCCTCCATGTCTGGGCTCTTCCCTGTCAACGCCGAGGCCCTCAAGTCTGTCTCTGGTG 1440
 Qy 481 ValLeuProLeuProAlaCysLeuAsnProLeuLeuTyrIleLeuPheAsnProHisPhe 500
 Db 1441 GTGCTGCCCTTGCTGCTGCCCTCAACCCACTGCTGTACCTGTCTTCAACCCCACTTC 1500
 Qy 501 ArgAspAspLeuArgArgLeuArgProArgAlaGlyAspSerGlyProLeuAlaTyrAla 520
 Db 1501 CGGGATGACCTTCCGGGGCTTCGGCCCCCGCGAGGGAGCTCAGGGCCCCCTAGCTATGCT 1560
 Qy 521 AlaAlaGlyGluLeuGluLysSerSerCysAspSerThrGlnAlaLeuValAlaPheSer 540
 Db 1561 GCGGCCGGGAGCTGGAGAAGAGCTCTCTGTGATTTCTACCCAGGCCCTGTAGCTTCTCT 1620
 Qy 541 AspValAspLeuIleLeuGluAlaSerGluAlaGlyArgProProGlyLeuGluThrTyr 560
 Db 1621 GATGTGATCTCATTTCTGGAAGCTTCTGAAGCTGGCGGGCCCCCTGGGGCTGGAGACCTAT 1680
 Qy 561 GlyPheProSerValThrLeuIleSerCysGlnGlnProGlyAlaProArgLeuGluGly 580
 Db 1681 GGCTTCCCCCTCAGTGACCTCTATCTCTGTGACGAGCAGGGGGCCCCCAGCTGGAGGGC 1740
 Qy 581 SerHisCysValGluProGluGlyAsnHisPheGlyAsnProGlnProSerMetAspGly 600
 Db 1741 AGCCATTGTAGAGCCAGAGGGGAACCACTTTTGGGAACCCCCCAACCTCCATGATGGA 1800
 Qy 601 GluLeuLeuLeuArgAlaGluGlySerThrProAlaGlyGlyLeuSerGlyGlyGly 620
 Db 1801 GAACGTGCTGTAGGGCAGAGGGATCTTACGCCAGAGCTGGAGGCTTGTGAGGGGTGGC 1860

```
Qy 621 GlyPheGlnProSerGlyLeuAlaPheAlaSerHisVal 633
Db 1861 GGCTTTACGCCCTCTGGCTTGGCTTTGCTTTCACACGTG 1899

RESULT 2
AX301825 2486 bp DNA linear PAT 30-NOV-2001
LOCUS AX301825
DEFINITION Sequence 4 from Patent WO0185768.
ACCESSION AX301825
VERSION AX301825.1 GI:17382883
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Homo sapiens
REFERENCE 1
AUTHORS Gu, W.
TITLE Novel g-protein coupled receptors and uses therefor
JOURNAL Patent: WO 0185768-A 4 15-NOV-2001;
MILLennium Pharmaceuticals, Inc. (US)
FEATURES
source
1. 2486
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
<2..1903
/notes="unnamed protein product"
/codon_start=1
/protein_id="CAD13077.1"
/db_xref="GI:17382884"
translations="NTTHYRESWACYRYSRIGPSTHASVERSGQLSLPAPASLAAL
AASNTASGKLELFTFQSLQALDLSWNAIRSHPEAFSTLHSLVKLDLTDNLQTT
LPLAGLGLHLKLNLAISQAFSKDPSFKRLILEVPYAYQCCPYGMCASFKKASGP
WEADLHLDDESSKRLPLGLLAQEAENHYDDLDLELMEDSKPHSPVQCSPTPGPF
KPCYLVESGIRLAWAI VLLSVLCNGLVLLTVFAGGPAPLPVPKFWVGAAGANTL
TGISGLASVDLTFQSEYGRWETGLGCRATGLAVLGSEASVLLLTAAVQCS
VSVSCRAYGSLSPSVRAGVLGCLAGLAAALPLASGEYASPLCLPYAPPEGQ
PALGFVTFALVMNNSFCFLVWAGYIKLYCDLPRGDFEAVWDCAMVRHVAWLI FADGL
LYCVAFLFSAFMGLFPVTPPEAVKSVLLVLLPLACINPLLLYLLFNPHFRDRLRLR
PRAGDSGLPAVAAGELKSSCDTQALVAPSDVDLILEASEAGRPGLFTYGPSPVT
LISQQPGAPLESCHVEPEGNHFGNPPQPSMDGELLRLRBAGSTPAGGGLSGGGFPQ
SGLAFASHV"

Alignment Scores:
Pred. No.: 2,25e-226 Length: 2486
Score: 3305.00 Matches: 632
Percent Similarity: 99.84% Conservative: 0
Best Local Similarity: 99.84% Mismatches: 1
Query Match: 99.94% Indels: 0
DB: 6 Gaps: 0

US-10-664-667-5 (1-633) x AX301825 (1-2486)

Qy 1 AsnThrThrHisTyrArgGluSerTrpTyrAlaCysArgTyrArgSerGlyIleProGly 20
Db 2 AATACGACTACTATAGGAAAGCTGTGACCCCTGCAGGTACCGCTCGGAATTCCTCCGGG 61

Qy 21 SerThrHisAlaSerValGluArgSerGlnGlyLeuSerLeuProAlaHisProAlaSer 40
Db 62 TCGACCCACGGCTCCGTGGAGCGGAGCAGGGCTGTGAGCCCTGCGGGCTCATCCAGCCCTCT 121

Qy 41 LeuAlaAlaLeuAlaAlaSerAsnThrThrAlaSerGlyLysLeuGlu**AspThrPhe 60
Db 122 CTTGCTGCCCTTAGCGGCTCCACACACACCCGATCTGGGAAATTGGAGCTNGACACCTTC 181

Qy 61 SerGlnLeuSerSerLeuGlnAlaLeuAspLeuSerTrpAsnAlaIleArgSerIleHis 80
Db 182 AGCCAGCTGAGCTCCCTCGCAAGCCCTCGATCTTAGCTGGAACGCCATCCGGTCCATCCAC 241

Qy 81 ProGluAlaPheSerThrLeuHisSerLeuValLysLeuAspLeuThrAspAsnGlnLeu 100
Db 242 CCTGAGGCCCTTCTCCACCCCTGACTCCCTGGTCAAGCTGGACCTGGACAGAACCCAGCTG 301
```

ORIGIN

```
Qy 101 ThrThrLeuProLeuAlaGlyLeuGlyLeuMetHisLeuLysLeuLysGlyAsnLeu 120
Db 302 ACCACACTGCCCTTGGCTGGACTTTGGGGCTTGTGTCATCTGAAGCTCAAGGGAACCTT 361

Qy 121 AlaLeuSerGlnAlaPheSerLysAspSerPheProLysLeuArgIleLeuGluValPro 140
Db 362 GCTCTCTCCCAAGGCTTCTTCCCAAGACAGTATTCCCAAACTGAGGATCCTCGAGGTGCT 421

Qy 141 TyrAlaTyrGlnCysCysProTyrGlyMetCysAlaSerPhePheLeuAlaSerGlyGln 160
Db 422 TATGCCTACCAAGTCTGCTCCCTATGGGATGTGCCAGCTTCTTCAAGGCCCTCTGGGGAG 481

Qy 161 TrpGluAlaGluAspLeuHisLeuAspAspGluGluSerSerLysArgProLeuGlyLeu 180
Db 482 TGGGAGGCTGAAGACCTTTCACCTTGTATGATGAGGAGTCTTCAAAAAGGCCCTTGGGCCTC 541

Qy 181 LeuAlaArgGlnAlaGluAsnHisTyrAspGlnAspLeuAspGluLeuGlnLeuGluMet 200
Db 542 CTTGCCACAGACAGCAGAGAACCACTATGACAGACCTGGATGAGCTCCAGCTGGAGATG 601

Qy 201 GluAspSerLysProHisProSerValGlnCysSerProThrProGlyProPheLysPro 220
Db 602 GAGGACTCAAAGGCCACACCCCGTGTCCAGTGTAGCCCTACTCCAGGCCCTTCAAGGCC 661

Qy 221 CysGluTyrLeuPheGluSerTrpGlyIleArgLeuAlaValTrpAlaIleValLeuLeu 240
Db 662 TGTGAGTACCTCTTTGAAAGCTGGGGCATCCGCTGGCCGTGGGCCATCGTGTGTGCTC 721

Qy 241 SerValLeuCysAsnGlyLeuValLeuLeuThrValPheAlaGlyGlyProAlaProLeu 260
Db 722 TCCGTGCTCTGCAATGAGCTGGTGTGTCGACCGTGTTCGTGGCGGCCCTGCCCCCTG 781

Qy 261 ProProValLysPheValValGlyValAlaIleAlaGlyAlaAsnThrLeuThrGlyIleSer 280
Db 782 CCCCCGTTCAGTTTGTGTAGTGGCATTCGAGCGCCCAACACTTGTACCTGAGCGCATTTCC 841

Qy 281 CysGlyLeuLeuAlaSerValAspAlaLeuThrPheGlyGlnPheSerSerGlyTyrGlyAla 300
Db 842 TGTGSCCTCTTAGCCTCAGTCGATGCCCTGACCTTTGGTCAGTCTCTCTGAGTACGAGGCC 901

Qy 301 ArgTrpGluThrGlyLeuGlyCysArgAlaThrGlyPheLeuAlaValLeuGlySerGlu 320
Db 902 CGCTGGGAGACGGGGCTAGGCTGCGGGCCACTGGCTTCTGGCAGTAGTACTTTGGGTGCGAG 961

Qy 321 AlaSerValLeuLeuLeuThrLeuAlaAlaValGlnCysSerValSerValSerCysVal 340
Db 962 GCATCGGTGCTGTGCTCACTCTGCGCGCAGTGGCAGTGCAGCGTCTCCGTCTCCTGTGTCT 1021

Qy 341 ArgAlaTyrGlyLysSerProSerLeuGlySerValArgAlaGlyValLeuGlyCysLeu 360
Db 1022 CGGGCCTATGGGAAGTCCCCCTCCCTGGGCAGCGTTCGAGCAGGGGTCTTAGGCTGCCTG 1081

Qy 361 AlaLeuAlaGlyLeuAlaAlaLeuProLeuAlaSerValGlyGlyTyrGlyAlaSer 380
Db 1082 GCATGGCAGGGCTGGCGCGCCACTGCCCCCTGAGTGGGAGATAAGCGGGCTCC 1141

Qy 381 ProLeuCysLeuProTyrAlaProProGluGlyGlnProAlaAlaLeuGlyPheThrVal 400
Db 1142 CCATCTGCTGCTGCTTACGCGCCACCTGAGGGTCAAGCCAGCAGCGGCTTGGGCTTCCCGGTG 1201

Qy 401 AlaLeuValMetMetAsnSerPheCysPheLeuValValAlaGlyAlaTyrIleLysLeu 420
Db 1202 GCCCTGGTGATGATGAATCCTTCTGTTCCTGGTGGCGGGTGGCTTACATACTCAACTG 1261

Qy 421 TyrCysAspLeuProArgGlyAspPheGluAlaValTrpAspCysAlaMetValArgHis 440
Db 1262 TACTGTGACCTGCGCGGGCGCACTTTGAGGCCGTGTGGGAGCTGCGCATGTGTGAGGCAC 1321

Qy 441 ValAlaTrpLeuIlePheAlaAspGlyLeuLeuTyrCysProValAlaPheLeuSerPhe 460
Db 1322 GTGGCCCTGGCTCATCTTTCGACAGACGGGCTCCTCTACTGTTCCTGGTGGCCCTCTCAGCTTC 1381
```

```
Qy 461 AlaSerMetLeuGlyLeuPheProValThrProGluAlaValLysSerValLeuVal 480
Db 1382 GCCTCCATGCTGGGCTCTTCCCTGTACAGCCCGAGGGCGCTCAAGTCTGTCTCTGGTG 1441
Qy 481 ValLeuProLeuProAlaCysLeuAsnProLeuLeuTyrLeuPheAsnProHisPhe 500
Db 1442 GTGCTGCCCTGCCCTGCCTCAACCACTGCTGTACTCTTCAACCCGCCACTTC 1501
Qy 501 ArgAspLeuArgAlaGlyLeuArgProHrgAlaGlyAspSerGlyProLeuAlaTyrAla 520
Db 1502 CGGATGACCTTCGGCGGCTTCGGCCCGCGAGGGGAGCTCAGGGCCCTAGCCTATGCT 1561
Qy 521 AlaAlaGlyLeuLeuGlyLysSerSerCysAspSerThrGlnAlaLeuValAlaPheSer 540
Db 1562 CGGCGCGGGAGCTGGAGAGAGTCTCTGTGATTTACCCAGGCGCTGTAGCTTCTCT 1621
Qy 541 AspValAspLeuLeuLeuGluAlaSerGluAlaGlyArgProProGlyLeuGluThrTyr 560
Db 1622 GATGTGGATCTCATTTCTGAAGCTTCTGAAGCTGGCGGCCCTCGGCTGGAGCCTAT 1681
Qy 561 GlyPheProSerValThrLeuLeuSerCysGlnGlnProGlyAlaProArgLeuGluGly 580
Db 1682 GGCTTCCCTCAGTGACCTCATCTCTGTGTAGAGCCAGGGGCGCCCGAGGCTGGAGGC 1741
Qy 581 SerHisCysValGluProGluGlyAsnHisPheGlyAsnProGlnProSerMetAspGly 600
Db 1742 AGCCATTGTGTAGAGCCAGAGGGAGACCATTTGGGAACCCCAACCTCCATGTGATGGA 1801
Qy 601 GluLeuLeuLeuArgAlaGluGlySerThrProAlaGlyGlyLeuSerGlyGlyGly 620
Db 1802 GAACCTGCTGTAGGGCAGAGGATCTACGCCAGCAGGGGTGGAGCTTGTACAGGGGTGGC 1861
Qy 621 GlyPheGlnProSerGlyLeuAlaPheAlaSerHisVal 633
Db 1862 GGCTTTACAGCCCTCGGCTGGCCCTTTGCTTACACAGT 1900

RESULT 3
AX098217 LOCUS AX098217 2322 bp DNA linear PAT 30-MAR-2001
DEFINITION Sequence 129 from Patent WO0118542.
ACCESSION AX098217
VERSION AX098217.1 GI:13515330
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Homo sapiens
REFERENCE 1
AUTHORS Lee,J., Thompson,P. and Lillie,J.
TITLE Identification, assessment, prevention, and therapy of ovarian cancer
JOURNAL Cancer
PATENT Patent: WO 0118542-A 129 15-MAR-2001;
MILLNNIUM Millennium Predictive Medicine, Inc. (US)
FEATURES
source
1..2322
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

ORIGIN

Alignment Scores:
Pred. No.: 2,43e-215 Length: 2322
Score: 3150.00 Matches: 609
Percent Similarity: 99.84% Conservative: 1
Best Local Similarity: 99.67% Mismatches: 1
Query Match: 95.25% Indels: 1
DB: 6 Gaps: 0

US-10-664-667-5 (1-633) x AX098217 (1-2322)

Qy 23 HisAlaSerValGluArgSerGlnGlyLeuSerLeuProAlaHisProAlaSerLeuAla 42
Db 2 CACGCGTCCGTGGAGCGGAGCAGGGTCTGAGCCTGCGGGCTCATCCAGGCTCTCTTGCT 61
```

```
Qy 43 AlaLeuAlaAlaSerAsnThrThrAlaSerGlyLysLeuGlu***AspThrPheSerGln 62
Db 62 GCCCTAGCGGCTCCAAACAACCGCATCTGGGAAATTTGGAGCT-GACACCTTCAGCCAG 120
Qy 63 LeuSerSerLeuGlnAlaLeuAspLeuSerTrpAsnAlaIleArgSerIleHisProGlu 82
Db 121 CTGAGCTCCCTGCAAGCCCTGGATCTTAGCTGGAAACGCCATCCGGTCCATCCACCCGTAG 180
Qy 83 AlaPheSerThrLeuHisSerLeuValLysLeuAspLeuThrAspAsnGlnLeuThrThr 102
Db 181 GCCTTCTCCACCTCGCATCTCCCTGTCAGCTGACCTGACAGACAAACCACTGACCACA 240
Qy 103 LeuProLeuAlaGlyLeuGlyGlyLeuMetHisLeuLysLeuLysGlyAsnLeuAlaLeu 122
Db 241 CTGCCCCCTGGCTGACTTGGGGGCTTGATGATCTGAAGCTCAAAAGGAACCTTGTCTCTC 300
Qy 123 SerGlnAlaPheSerLysAspSerPheProLysLeuArgIleLeuGluValProTyrAla 142
Db 301 TCCAGGCGCTTCTCCAAGGACAGTTTCCCAAACTGAGGATCCTGAAGGTGCCTTATGCC 360
Qy 143 TyrGlnCysCysProTyrGlyMetCysAlaSerPhePheLysAlaSerGlyGlnTrpGlu 162
Db 361 TACCAGTGTCTCTCCCTATGGGATGTGTGCCAGCTTCTTCAAGGCTCTTGGGACAGTGGAG 420
Qy 163 AlaGluAspLeuHisLeuAspAspGluLysSerSerLysArgProLeuGlyLeuAlaLeu 182
Db 421 GCTGAAGACCTTCACTTGTATGATGAGAGGTCTTCAAAAGGGCCCTTGGGCTCTTGTGCC 480
Qy 183 ArgGlnAlaGluAsnHisTyrAspGlnAspLeuAspGluLeuGlnLeuGluMetGluAsp 202
Db 481 AGACAGCAGAGAACCACTATGACCAGGACCTGGATGAGCTCCAGCTGGAGATGGAGGAC 540
Qy 203 SerLysProHisProSerValGlnCysSerProThrProGlyProPheLysProCysGlu 222
Db 541 TCAAAGCCACACGCCCTGCTCAGTGTAGGCCCTACTCCAGGCCCTTCAAGCCCTGTGAG 600
Qy 223 TyrLeuPheGluSerTrpGlyIleArgLeuAlaValTrpAlaIleValLeuLeuSerVal 242
Db 601 TACTCTTTTGAAGCTGGGGCATCCGCTGGCGTGGGGCCATCGTGTGTCTCTCCGTG 660
Qy 243 LeuCysAsnGlyLeuValLeuLeuLeuThrValPheAlaGlyGlyProAlaProLeuProPro 262
Db 661 CTCTGCAATGGACTGGTGTGCTGACCGTGTCTGCTGGCGGCCCTGCCCCCTGCCCCG 720
Qy 263 ValLysPheValValGlyAlaIleAlaGlyAlaAsnThrLeuThrGlyIleSerCysGly 282
Db 721 GTCAAGTTTGTGGTAGGTGCGATTGTCAGGGCGCCAAACACCTTGACTGGCATTTCTGTGGC 780
Qy 283 LeuLeuAlaSerValAspAlaLeuThrPheGlyGlnPheSerGluTyrGlyAlaArgTrp 302
Db 781 CTCTAGCCCTCAGTCGATGCCCTGACCTTTGGTCAGTTCCTCTGAGTAGGAGGCCCTGTGG 840
Qy 303 GluThrGlyLeuGlyCysArgAlaThrGlyPheLeuAlaValLeuGlySerGluAlaSer 322
Db 841 GAGCGGGGTAGGCTGCCGGGCCACTGGCTTCTTGGCAGTACTTGGGTGGAGGCATCG 900
Qy 323 ValLeuLeuLeuThrLeuAlaValGlnCysSerValSerValSerCysValArgAla 342
Db 901 GTGCTGTGCTCACTCTGGCCCGCAGTGCAGTGCAGCGCTCTCCGTCTCTGTGTCCGGGCC 960
Qy 343 TyrGlyLysSerProSerLeuGlySerValArgAlaGlyValLeuLeuCysLeuAlaLeu 362
Db 961 TATGGGAAAGTCCCCCTCCCTGGGCGAGGTTCTGAGCAGGGGTCTCTAGGCTGCCTGGCACTG 1020
Qy 363 AlaGlyLeuAlaAlaLeuProLeuAlaSerValGlyGluTyrGlyAlaSerProLeu 382
Db 1021 GCAGGGCTGGCCGCGCAGCTGCCCTGGCCCTCAGTGGGAGAAATACGGGGCTCTCCCACTC 1080
Qy 383 CysLeuProTyrAlaProProGluGlyGlnProAlaAlaLeuGlyPheThrValAlaLeu 402
Db 1081 TGCCTGCGCTTACGCGGCCACCTGAGGGTTCAGCCAGCAGCCCTGGGCTTTCACCGTGGCCCTG 1140
```



```
Db 1794 GTAGGTGGATTGAGGGCCCAACACCTTGACTGGCATTTCTGTGGCTTCTAGCCCTCA 1853
Qy 287 ValAspAlaLeuThrPheGlyGlnPheSerGluTyrGlyAlaAraGfTrpGluThrGlyLeu 306
Db 1854 GTCGATGCCCTGACCTTTGGTCAGTTCTCTGAGTACGGAGCCGCTGGAGACGGGGCTA 1913
Qy 307 GlyCysArgAlaThrGlyPheLeuAlaValLeuGlySerGluAlaSerValLeuLeuLeu 326
Db 1914 GGCTGCGCGGCCACTGGCTTCTGGCAGTACTTGGGTGGAGGCATCGGTGCTGCTGCTC 1973
Qy 327 ThrLeuAlaAlaValGlnCysSerValSerCysValArgAlaTyrGlyLysSer 346
Db 1974 ACTTGGCGCGAGTGCAGTGCAGCGTCTCCGTCTCTGCTGTCGGGGCTATGGGAGTCC 2033
Qy 347 ProSerLeuGlySerValArgAlaGlyValLeuGlyCysLeuAlaLeuAlaGlyLeuAla 366
Db 2034 CCTTCCCTGGGCAGCGTTCGAGCAGGGGCTCTAGGCTGCCTGGCACGTGGCAGGGCTGCC 2093
Qy 367 AlaAlaLeuProLeuAlaSerValGlyGluTyrGlyAlaSerProLeuCysLeuProTyr 386
Db 2094 GCCGACCTGCCCTTGGGCTCAGTGGGAGAAATACGGGGCTTCCCACTCTGCTGCCCTAC 2153
Qy 387 AlaProGluGlyGlnProAlaAlaLeuGlyPheThrValAlaLeuValMetMetAsn 406
Db 2154 GCGCACCTGAGGTCAGCCAGCAGCCCTGGGCTTCCACGTGGGCTTGGTATGATGAAC 2213
Qy 407 SerPheCysPheLeuValAlaGlyAlaTyrIleLysLeuTyrCysAspLeuProArg 426
Db 2214 TCCCTTCTGTTTCCCTGGTGGTGGGCGGCTGACATCAAACTGATCTGTGACCTGCCCGG 2273
Qy 427 GlyAspPheGluAlaValTrpAspCysAlaMetValArgHisValAlaTrpLeuIlePhe 446
Db 2274 GGCCACTTGGGCGCGTGTGGGACTGCGGCCATGTGAGGCACGTGGGCTGGCTCATCTTC 2333
Qy 447 AlaAspGlyLeuLeuTyrCysProValAlaPheLeuSerPheAlaSerMetLeuGlyLeu 466
Db 2334 GCAGCGGGCTCTCTACTGTCGCGGCTTCTCTCAGCTTCGGCTCCATGCTGGGCTC 2393
Qy 467 PheProValThrProGluAlaValLysSerValLeuLeuValValLeuProLeuProAla 486
Db 2394 TTCCCTGTCAAGCGCGGCGGCTCAAGTCTGTCTGCTGCTGGTGTGCTGCCCTGCCCTGCC 2453
Qy 487 CysLeuAsnProLeuLeuTyrLeuLeuPheAsnProHisPheArgAspAspLeuArgArg 506
Db 2454 TGCCTCAACCCACCTGCTGTACTGCTCTTCAACCCCACTTCCGGGATGACCTTTCGGCGG 2513
Qy 507 LeuArgProArgAlaGlyAspSerGlyProLeuAlaTyrAlaAlaAlaGlyGluLeuGlu 526
Db 2514 CTTGCGCCCGCGCAGGGGACTCAGGGCCCTAGCCTATGCTGCGCGCGGAGCTGGAG 2573
Qy 527 LysSerSerCysAspSerThrGlnAlaLeuValAlaPheSerAspValAspLeuLeuLeu 546
Db 2574 AAGAGCTCCTGTGATCTACCCAGGCGCTGGTACCTTCTCTGATGTGGATCTCATTCGTG 2633
Qy 547 GluAlaSerGluAlaGlyArgProProGlyLeuGluThrTyrGlyPheProSerValThr 566
Db 2634 GAAGCTTCTGAAGCTGGCGCGCCCTTGGGCTGGAGACCTATGGCTTCCCTCAGTGACC 2693
Qy 567 LeuIleSerCysGlnProGlyAlaProArgLeuGluGlySerHisCysValGluPro 586
Db 2694 CTCACTCTCTGTAGCAGCCGCGGGGCCCCCGGCTGGAGGGCAGCCATTTGTAGAGCCA 2753
Qy 587 GluGlyAsnHisPheGlyAsnProGlnProSerMetAspGlyGluLeuLeuLeuArgAla 606
Db 2754 GAGGGGAACCACTTTGGGAACCCCAACCTCCATGATGGAGAACTGCTGCTGAGGGCA 2813
Qy 607 GluGlySerThrProAlaGlyGlyLeuSerGlyGlyGlyPheGlnProSerGly 626
Db 2814 GAGGGAATCTACGCAGGAGGTGGAGCTTGTGAGGGGGTGGCGGCTTTCAGCCCTCTGCC 2873
Qy 627 LeuAlaPheAlaSerHisVal 633
```

```
Db 2874 TTGGCCTTTGCTTACACACGTG 2894
RESULT 5
AX301830
LOCUS
DEFINITION Sequence 9 from Patent WO0185768.
ACCESSION AX301830
VERSION AX301830.1 GI:17382889
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1
Gu, W.
REFERENCE
TITLE Novel g-protein coupled receptors and uses therefor
JOURNAL Patent: WO 0185768-A 9 15-NOV-2001;
Millennium Pharmaceuticals, Inc. (US)
FEATURES
Location/Qualifiers
1..2208
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
<1..>2208
/notes="unnamed protein product"
/codon_start=1
/protein_id="CAD13080.1"
/db_xref="GI:17382890"
/translating="GLHNETLDLNVNKLQEPFVAITLQRLQELGFHNNNIKAPEK
AFMGNPLQTIHFYDNP IQVGRSAFOYLKPLHTLSLNGAMDIQEPDLKGTTSLEIL
TLFRAGIRLLPQGMCOQLPRLVLELSHNOIEELPSLHRCQKLEELGLQHNRIWEIGA
DTFSQSSLOALDLSMNAIRSIHPEAFSTLSLSLVKLDLTQNLTLTLAGLGGIMHLK
LKNLALSAQFSKDSFPKRLILEVPYAYCCPYGMCASFPAKSGQWAEADLHDDSES
SKRPLGLLAKAENHYDDLDLEQLEMEDSKPHSPVQCSPTGPFKCEYLFPESWGR
LAYMAVLSVLCNGLVLTVPAGGPAPLPVVFVGAIAAGATLTGISCGLLASVDA
LTGQSEYGARWETGRCATGFLASVGEASVLLITLAQVCSVSVSCVRYAKGSP
SLGSRVAGVGLGALAGLAALPLASVGEYASPLCLPYAPPEGQALGFTVALVM
NSFCFLVAVAGATIKLYCDLPRGDPEAVMDCAMVRHVAVLIPADGLLYCPVAFLSFAM
LGLFPTPEAKSVLLVPLPLACLPLLLYLFNPHRDLRLRLPRAGSGPLAYAA
AGLEKSSCDTQALVAFSDVLDILEASEAGRPGLETYGYFPSTVLLISCOQPGAPRL
GSHCVPEFGNHFGNPQPSMDGELLRAEGSTPAGGLSGGGGFPQPSGLAFASHV"
ORIGIN
Alignment Scores:
Pred. No.: 6,62e-208 Length: 2208
Score: 3045.50 Matches: 602
Percent Similarity: 92.61% Conservative: 12
Best Local Similarity: 90.80% Mismatches: 19
Query Match: 92.09% Indels: 32
DB: 6 Gaps: 4
US-10-664-667-5 (1-633) x AX301830 (1-2208)
Qy 1 AsnThrThrHisTyr-----
Db 224 AACTCCACACACTATCTCTGAATGGTCCATCCAGGAGTTTCCAGATCTCAAG 283
Qy 6 ---ArgGluSerTrpTyrAla-----CysArgTyrArg 15
Db 284 GCACCACAGCTGGAGATCTGTACCTGACCCCGCGCAGGCATCCGGCTGCTCCCATCGG 343
Qy 16 -----SerGlyIleProGlySer-ThrHisAlaSerValGluArgSerGln---G1 31
Db 344 GGATGTCGCAACAGCTGCCAGGCTCGAGTCTCTGGAACTGTCTCAACATCAAAATTGAGG 403
Qy 31 YLeuSerLeuProAlaHisProAlaSerLeuAlaAlaLeuAlaAlaSerAsnThrThrAl 51
Db 404 AGTGCCCGCAGCTGCACAGGTGTGAGAAATTGGAGGAA--TCGGCCTCCAACACACGCG 462
Qy 51 aSerGlyLysLeuGlu***AspThrPheSerGlnLeuSerSerLeuGlnAlaLeuAspLe 71
Db 463 ATCTGGGAATTTGGAGCT-GACACCTTCAGCCAGCTGAGCTCCCTGCAAGCCCTGGATCT 521
```


Qy 71 uSerTrpAlaAlaIleArgSerIleHisProGluAlaPheSerThrLeuHisSerLeuVa 91
Db 522 TAGCTGGAAACGCATCCGCTCCATCCACCTGAGGCTTCTCCACCTGCATCCCTGGT 581
Qy 91 llyLeuAspLeuThrAspAsnGlnLeuThrLeuProLeuAlaGlyLeuGlyLe 111
Db 582 CAAGCTGACCTTGACAGAACACCTGACACACTGCCCTTGGCTGGACTTGGGGGCTT 641
Qy 111 uMetHisLeuLeuLeuGlyAsnLeuAlaLeuSerGlnAlaPheSerLysAspSerPh 131
Db 642 GATGCATCTGAAGCTCAAGGGAACCTTGTCTCTCCAGGCTTCTCCAGGACAGTTT 701
Qy 131 eProLysLeuArgIleLeuGluValProTyrAlaTyrGlnCysCysProTyrGlyMetCy 151
Db 702 CCCAAACTGAGGATCTCGAGGTGCCTTATGCTTACCACTGCTGCTCCCTATGGATGTG 761
Qy 151 eAlaSerPhePheLeuAlaSerGlyGlnTrpGluAlaGluAspLeuHisLeuAspAspG1 171
Db 762 TGCCAGCTTCTTCAAGGCTCTGGGCAGTGGAGGCTGAAGACCTTCACTTGTATGATGA 821
Qy 171 uGluSerSerLysArgProLeuGlyLeuLeuAlaArgGlnAlaGluAsnHisTyrAspG1 191
Db 822 GGAGCTTCTCAAAAGGCCCTGGGCTCTTGGCCAGACAGCAGAACCTATGACCA 881
Qy 191 nAspLeuAspGluLeuGlnLeuGluMetGluAspSerLysProHisProSerValGlnCy 211
Db 882 GGACCTGATGAGCTCCAGCTGGAGTGGAGGACTCAAAAGGCACACCCAGTGTCCAGTG 941
Qy 211 eSerProThrProGlyProPheLysProCysGluTyrLeuPheGluSerTrpGlyIleAr 231
Db 942 TAGCCCTACTCCAGGCCCTTCAAGCCCTGTAGTACCTCTTTGAAAGCTGGGGATCCG 1001
Qy 231 gLeuAlaValTrpAlaIleValLeuLeuSerValLeuCysAsnGlyLeuValLeuLeuTh 251
Db 1002 CCTGGCCGTGGGCCCATCGTGTCTCTCGTGTCTTGCATGAGCTGGTGTCTGCTGAC 1061
Qy 251 rValPheAlaGlyGlyProAlaProLeuProProValLysPheValValGlyAlaIleAl 271
Db 1062 CGTGTTCCTGGCGGCCCTGCCCCCTGCCCGGCTCAAGTTTGTGTAGTGGCATTCG 1121
Qy 271 aGlyAlaAsnThrLeuThrGlyIleSerCysGlyLeuLeuAlaSerValAspAlaLeuTh 291
Db 1122 AGGCCCAACACCTTGTAGTGGCATTTCTGTGGCCCTTCTAGCCTCAGTGCAGTCCCTGAC 1181
Qy 291 rPheGlyGlnPheSerGluTyrGlyAlaArgTrpGluThrGlyLeuGlyCysArgAlaTh 311
Db 1182 CTTTGGTCAGTCTCTGAGTACGAGCCCGCTGGAGAGCGGGCTAGGCTGCCGGGCCAC 1241
Qy 311 rGlyPheLeuAlaValLeuGlySerGluAlaSerValLeuLeuLeuThrLeuAlaAlaVa 331
Db 1242 TGGCTTCTGGCAGTACTTGGTGGAGGCATCGGTGCTGTGCTCACTCTGGCGGCAGT 1301
Qy 331 lGlnCysSerValSerValSerCysValArgAlaTyrGlyLysSerProSerLeuGlySe 351
Db 1302 GCAGTGCAGCTTCCGCTCTCTGTGTCCGGCCCTATGGAGTCCCTCCCTGGGGAG 1361
Qy 351 rValArgAlaGlyValLeuGlyCysLeuAlaLeuAlaGlyLeuAlaAlaLeuProLe 371
Db 1362 GTTTCGACAGGGGTCCTAGGCTGCTGGCACTGGCAGGCTGGCCCGCCGACCTGCCCT 1421
Qy 371 uAlaSerValGlyGluTyrGlyAlaSerProLeuCysLeuProTyrAlaProProGluG1 391
Db 1422 GGCCTCAGTGGAGAAATACGGGCCCTCCCCACTCTGCTGCCCTACGCGCCACCTGAGGG 1481
Qy 391 yGlnProAlaAlaLeuGlyPheThrValAlaLeuValMetMetAsnSerPheCysPheLe 411
Db 1482 TCAGCCACAGCCCTGGGCTTCCACGCTGGCCCTTGTGTATGATGAACCTCTTCTGTTCCT 1541
Qy 411 uValValAlaGlyAlaTyrIleLeuLeuTyrCysAspLeuProArgGlyAspPheGluAl 431
Db 1542 GGTGTGCCCGGTGCTCATCAAACTGCTGTGACCTGCCCGGGCGGCACTTTGAGGC 1601
Qy 431 aValTrpAspCysAlaMetValArgHisValAlaTrpLeuIlePheAlaAspGlyLeuLe 451

Db 1602 CGTGTGGACTGCGCCATGTGTAGGCACAGTGGCCCTGGCTCATCTTCGACAGAGGGCTCCT 1661
Qy 451 uTyrCysProValAlaPheLeuSerPheAlaSerMetLeuGlyLeuPheProValThrPr 471
Db 1662 CTACTGTCCCGTGGCCCTTCTCAGCTTCCGCTTCCATGCTGGGCCCTTCTTCCCTGTACGCC 1721
Qy 471 oGluAlaValLysSerValLeuLeuValValLeuProLeuProAlaCysLeuAsnProLe 491
Db 1722 CGAGGCGGTCAAGTCTGTCTGTGTGGTGTCTGCCCTGTGCTGCTGCCCTCAACCCACT 1781
Qy 491 uLeuTyrLeuLeuPheAsnProHisPheArgAspLeuArgLeuArgLeuArgProArgAl 511
Db 1782 GCTGTACTGTCTTCAACCCCACTTCCGGGATGACCTTCGGCGGCTTCGGCCCCGGGC 1841
Qy 511 aGlyAspSerGlyProLeuAlaTyrAlaAlaGlyGluLeuGlyLysSerSerCysAs 531
Db 1842 AGGGGACTCAGGGGCCCTTAGGCTATGTGCGCGCGGGAGCTGGAGAAGAGCTCTCTGTGA 1901
Qy 531 pSerThrGlnAlaLeuValAlaPheSerAspValAspLeuIleLeuGluAlaSerGluAl 551
Db 1902 TTCTACCCAGGCCCTGGTAGCTTCTCTGATGTGGATCTCATTTCTGGAAGCTTCTGAAGC 1961
Qy 551 aGlyArgProProGlyLeuGluThrTyrGlyPheProSerValThrLeuIleSerCysG1 571
Db 1962 TGGCGCGGCCCTGGGCTGGAGACTATGGCTTCCCTCAGTGACCTCATCTCCTGTCA 2021
Qy 571 nGlnProGlyAlaProArgLeuGlySerHisCysValGluProGluGlyAsnHisPh 591
Db 2022 GCAGCCAGGGGCCCTCAGGCTGGAGGCGAGCCATTGTGTAGCCAGAGGGGAACCACTT 2081
Qy 591 eGlyAsnProGlnProSerMetAspGlyGluLeuLeuLeuArgAlaGluGlySerThrPr 611
Db 2082 TGGGAACCCCAACCTCCATGGATGGAGAACTGCTGTGAGGCGAGAGGATCTTACGCC 2141
Qy 611 oAlaGlyGlyLeuSerGlyGlyGlyPheGlnProSerGlyLeuAlaPheAlaSe 631
Db 2142 AGCAGGTGGAGGCTTGTGAGGGGTGGCGGCTTTTCAGCCCTCTGGCTTGGCTTTC 2201
Qy 631 rHisVal 633
Db 2202 ACACGT 2208

RESULT 6

LOCUS AX301828 2711 bp DNA linear PAT 30-NOV-2001
DEFINITION Sequence 7 from Patent WO0185768.
ACCESSION AX301828
VERSION AX301828.1 GI:17382887
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE

1 Gu, W.
Novel g-protein coupled receptors and uses therefor
Patent: WO 0185768-A 7 15-NOV-2001;
Millennium Pharmaceuticals, Inc. (US)

FEATURES

source

1. .2711
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

CDS

1. .2211
/note="unnamed protein product"
/codon_start=1
/protein_id="CAD13079.1"
/db_xref="GI:17382888"

/translation="GLNLESLDNLNKLQEFVPAIRTLGRLQELGFHNNNIKAIPK
AFMGNPLQTIHFYDNPQFVGRSAFYLPKLHLSUNGAMDIQEPDLKGTTSLEIL
TLTRAGIRLLPSCMCQQLPRLRLSHNQIEELPSLHRCQKBEI GLQHNRIWEIGA
DTFSQLSLQALDLSMNAIRSIHPEAFSLHSLVKLDLTDNLQTLTLPLAGLGLMLHK

LKGNLALSOAFSKDSFPKLRILEVPYAYOCCPYGMCASFFKASQWBAEDLHLDERS
SKRPLGILLARENHYDDELQLEMEDSKPHSVQCSPTGPKKCEYLFEWSGIR
LAWAIVLLSVLCNGVLVLTVPAGPAPLPVKFVGAIGANTILGTISCGLLASVDR
LTFGQFSYGRWETGLGCRATGFLAVLGSEASVLLLTAAVQCSVSVRAYGKSP
SLGSVRAGVAGCLALAGLAALPLASVGEYGAELCLIPYAPPEQGPAAALGFTVALVM
NSFCFLVAGAYIKLYCDLPRGDEAVWDCAMVRHAWLIIFADGLLYCPVAPLSPASM
LGLFPVTPKAVKSLVLLVPLPACLNPLLYLLFNPHPRDLRLRRPRAGDSGLPAYAA
AGELEKSDSDSTQALVAFSDVDLILEASEAGRPGLGYTSPVTLISCOOPGAPRLE
GSHCVPEPNHFGNPQPSMDGELLURLABGSTPAGUGSGGGGFPQPSGLAFASHV"

ORIGIN

Alignment Scores:

Pred. No.: 8, 44e-208 Length: 2711
Score: 2045.50 Matches: 602
Percent Similarity: 92.61% Conservative: 12
Best Local Similarity: 90.80% Mismatches: 19
Query Match: 92.09% Indels: 32
DB: 6 Gaps: 4

US-10-664-667-5 (1-633) x AX301828 (1-2711)

Qy 1 AenThrHisTyr----- 5
Db 224 AACTCCACACACTATCTCTGAATGGTGCATGGACATCCAGAGTTTCCAGATCTCAAG 283
Qy 6 ---ArgGluSerTyrAla-----CysArgTyrArg 15
Db 284 GCACACAGCCTGGAGATCTGTACCCCTGACCCCGCAGGACATCCGGCTGCTCCATCGG 343
Qy 16 -----SerGlyIleProGlySer-ThrHisAlaSerValGluArgSerGln---G1 31
Db 344 GGATGTGCCAACAGCTGCCAGGCTCCAGTCTCGGAACCTGTCTCACAAATTTGAGG 403
Qy 31 YLeuSerLeuProLHisProAlaSerLeuAlaAlaLeuAlaAlaSerAenThrAl 51
Db 404 AGCTGCCACGCTGCACAGGTGTGAGAAATTTGGAGAAA-TCGGCCTCCAACACACCGC 462
Qy 51 aSerGlyIleLeuGlu***AspThrPheSerGlnLeuSerSerLeuGlnAlaLeuAspLe 71
Db 463 ATCTGGGAAATTTGGAGCT-GACACCTTCAGCCAGCTGAGCTCCCTGCAAGCCCTGGATCT 521
Qy 71 uSerTrpAsnAlaIleArgSerIleHisProGluAlaPheSerThrLeuHisSerLeuVa 91
Db 522 TAGCTGGNACGCCATCCGCTCCATCCACCTTGAGGCTTCTCCACCTTGCACTCCCTGGT 581
Qy 91 lLysLeuAspLeuThrAspAenGlnLeuThrThrLeuProLeuAlaGlyLeuGlyLe 111
Db 582 CAAGCTGGACCTGACAGAACACAGCTGACCACACTGCCCCCTGGCTGGACTTGGGGGCTT 641
Qy 111 uMetHisLeuIleuIleuIleuIleuIleuIleuIleuIleuIleuIleuIleuIleuIleu 131
Db 642 GATGCATCTGAAGCTCAAAAGGAACTTGTCTCTCCAGGCTTCTCCAAAGGACAGTTT 701
Qy 131 eProIleuArgIleLeuGluValProTyrAlaTyrGlnCysCysProTyrGlyMetCy 151
Db 702 CCCAAACTGAGGATCTCGAGGTGCCTTATGCCATACAGTGCCTGCTCCCTATGGATGTG 761
Qy 151 sAlaSerPhePheIleuAlaSerGlyGlnTrpGluAlaGluAspLeuHisLeuAspAspG1 171
Db 762 TGCCAGCTTCTTCAAGGCTCTGGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAG 821
Qy 171 uGluSerSerIleuIleuIleuIleuIleuIleuIleuIleuIleuIleuIleuIleuIleu 191
Db 822 GGAGTCTTTAAAAAGGCCCCCTGGGCTCTTGGCAGACACAGCAGAGAACCACTATGACCA 881
Qy 191 nAspLeuAspGluLeuGlnLeuGluMetGluAspSerIleuIleuIleuIleuIleuIleuIleu 211
Db 882 GGACTGGATGAGCTCCAGCTGGAGATGGAGATCTCAAGGCCACACCCCGAGTTCAGTG 941
Qy 211 sSerProThrProGlyProPheIleuIleuIleuIleuIleuIleuIleuIleuIleuIleuIleu 231
Db 942 TAGCCCTACTCCAGGCCCTTCAAGCCCTGTGAGTACTCTTTGAAAGCTGGGCAATCCG 1001

Qy 231 gLeuAlaValTroAlaIleValLeuLeuSerValLeuCysAsnGlyLeuValLeuLeuTh 251
Db 1002 CTTGGCCGTGGGCCATCGTGTCTCTCGTCTCTCGTCTCTCGTCTCTCGTCTCTCGTCTCT 1061
Qy 251 rValPheAlaGlyProAlaProLeuProProValLysPheValValGlyAlaIleAl 271
Db 1062 CGTGTCTGGTGGGGGCTGCCGCCCTGCCCGGTCAAGTTTGTGTGTGTGTGTGTGTGTGTGT 1121
Qy 271 aGlyAlaAsnThrLeuThrGlyIleSerCysGlyLeuLeuAlaSerValAspAlaLeuTh 291
Db 1122 AGGCCCAACACCTTGATGGCATTTCTGTGGCTTCTAGCCTCAGTCAGTCAGTCAGTCAG 1181
Qy 291 rPheGlyGlnPheSerGluTyrGlyAlaArgTyrGluThrGlyLeuGlyCysArgAlaTh 311
Db 1182 CTTTGGTCAGTCTCTGAGTACGAGCCCTGGAGACGGGTAGCTGCCGGGCCAC 1241
Qy 311 rGlyPheLeuAlaValLeuGlySerGluAlaSerValLeuLeuLeuLeuThrLeuAlaVa 331
Db 1242 TGGCTTCTTGGCAGTACTTGGGTGGAGGCATCGTGTCTGTCTGTCTGTCTGTCTGTCTGTCT 1301
Qy 331 lGlnCysSerValSerValSerCysValArgAlaTyrGlyIleSerProSerLeuGlySe 351
Db 1302 GCAGTGCAGGCTCTCCGTCTCTCGTGTCCGGGCTATGGGAAGTCCCTCCCTGGGCAG 1361
Qy 351 rValArgAlaGlyValLeuGlyCysLeuAlaLeuAlaGlyLeuAlaAlaLeuProLe 371
Db 1362 CGTTCGAGCAGGGGTCTTAGCTGCTTGGCAGTGGCAGGCTGGCCCGCCGACCTGCCCCCT 1421
Qy 371 uAlaSerValGlyIleGluTyrGlyAlaSerProLeuCysLeuProTyrAlaProGluG1 391
Db 1422 GGCTCAGTGGAGATACGGGGCTTCCCTCCTCTGCTTGGCTTGGCTTGGCTTGGCTTGGCT 1481
Qy 391 YGlnProAlaAlaLeuGlyPheThrValAlaLeuValMetMetAsnSerPheCysPheLe 411
Db 1482 TCAGCCAGCAGCCTGGGCTTCCCGTGGGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1541
Qy 411 uValValAlaGlyAlaTyrIleValLeuTyrCysAspLeuProArgGlyVAspPheGluAl 431
Db 1542 GGTGTGGCGGGTGTCTTACATCAAACTGTACTGTGACCTGCCGGGGGAGCTTTGAGGC 1601
Qy 431 aValTrpAspCysAlaMetValArgHisValAlaTrpLeuIlePheAlaAspGlyLeuLe 451
Db 1602 CGTGTGGACTTGGCCATGTGTGGGACAGCTGGGCTTGGCTTGGCTTGGCTTGGCTTGGCT 1661
Qy 451 uTyrCysProValAlaPheLeuSerPheAlaSerMetLeuGlyLeuPheProValThrPr 471
Db 1662 CTACTGTCCGTGGCTTCTCAGCTTGGCTTCCATGTCTGGGCTTCTTGTGTGTGTGTGTGT 1721
Qy 471 oGluAlaValLysSerValLeuValValLeuProLeuProAlaCysLeuAsnProLe 491
Db 1722 CGAGGCGCTCAAGTCTGTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1781
Qy 491 uLeuTyrLeuLeuPheAsnProHisPheArgAspAspLeuArgLeuArgProArgAl 511
Db 1782 GCTGTACTTCTTCAACCCCTTCCGGGATGACCTTCCGGGGCTTCCGGCCCCCGGC 1841
Qy 511 aGlyAspSerGlyProLeuAlaTyrAlaAlaGlyGluLeuGluLysSerSerCysAs 531
Db 1842 AGGGAGCTCAGGGGCTTATGCTGTGGGCGGGGAGCTGGAGAAGAGCTCTCTGTGA 1901
Qy 531 pSerThrGlnAlaLeuValAlaPheSerAspValAspLeuIleuGluAlaSerGluAl 551
Db 1902 TTCTACCGAGGCTGT 1961
Qy 551 aGlyArgProProGlyLeuGluThrTyrGlyPheProSerValThrLeuIleSerCysG1 571
Db 1962 TGGCGGCCCCCTGGGCTGGAGACCTATGGCTTCCCTCAGTGACCTCATCTCTCTCTCA 2021
Qy 571 nGlnProGlyAlaProArgLeuGluGlySerHisCysValGluProGluGlyAsnHisph 591
Db 2022 GCAGCCAGGGGCCCCCAGGCTGGAGGCGAGCCATTTGTGTAGAGCCAGAGGAGAACCACTT 2081
Qy 591 eGlyAsnProGlnProSerMetAspGlyGluLeuLeuArgAlaGluGlySerThrPr 611

Qy 371 uAlaSerValGlyGluTyrGlyAlaSerProLeuCysLeuProTyrAlaProProGluGl 391
 Db 2115 GGCCTCACTGGAGATACGGGGCTCCCACTCTGCTGCCCTACGGCCACCTGAGG 2174
 Qy 391 yGlnProAlaLeuGlyPheThrValAlaLeuValMetMetAenSerPheCysPheLe 411
 Db 2175 TCAGCCAGCAGCCCTGGGCTTCACGCTGGCCCTGGTGATGATGAACCTCTCTGTTCT 2234
 Qy 411 uValValAlaGlyAlaTyrTrilleLeuTyrCysAspLeuProArgGlyAspPheGluAl 431
 Db 2235 GGTGTGGCCGGTGCCTACATCAAACTGTACTGTGACCTGCCGGGGCGACTTTGAGGC 2294
 Qy 431 aValTyrAspCysAlaMetValArgHisValAlaTyrLeuLeuPheAlaAspGlyLeuLe 451
 Db 2295 CGTGTGGGACTGGCCCATGGTGAGGCACGTGGCTGGCTCATCTTCGACAGCGGCTCCT 2354
 Qy 451 uTyrCysProValAlaPheLeuSerPheAlaSerMetLeuGlyLeuPheProValThrPr 471
 Db 2355 CTACTGTCCCGTGGCTTCTCAGCTTCCGCTCCATGCTGGGCTCTTCCCTGTACGCC 2414
 Qy 471 oGluAlaVallySerValLeuLeuValValLeuProLeuProAlaCysLeuAenProLe 491
 Db 2415 CGAGCCCGTCAAGTCTGTCTGTGGTGGTGTCTGCCCTGGCTGGCTTCCCAACCCACT 2474
 Qy 491 uLeuTyrLeuLeuPheAenProHisPheArgAspLeuArgLeuArgProArgAl 511
 Db 2475 GCTGTACTGTCTTCAACCCCGCTTCGGGATGACCTTCGGGGCTTCGGCCCGCGC 2534
 Qy 511 aGlyAspSerGlyProLeuAlaTyrAlaAlaAlaGlyGluLeuGlySerSerCysAs 531
 Db 2535 AGGGGACTCAGGCGCCCTAGCTATGCTGCGCGCGGAGCTGGAGAAGACTCCTGTGA 2594
 Qy 531 pSerThrClnAlaLeuValAlaPheSerAspValAspLeuLeuAlaSerGluAl 551
 Db 2595 TTCACCCAGCCCTGGTAGCTTCTCTGATGTGGATCTCATCTCGGAAGCTTCTGAAGC 2654
 Qy 551 aGlyArgProGlyLeuGluThrTyrGlyPheProSerValThrLeuLeuSerCysGl 571
 Db 2655 TGGCGGGCCCTGGCTGGAGACTATGGCTTCCCTCAGTGACCTCATCTCTGTCA 2714
 Qy 571 nGlnProGlyAlaProArgLeuGlySerHisCysValGluProGluGlyAenHisPh 591
 Db 2715 GCAGCCAGGGGCCCCAGGCTGGAGGCGAGCCATTGTGTAGAGCCAGAGGGAACCACTT 2774
 Qy 591 eGlyAsnProGlnProSerMetAspGlyGluLeuLeuLeuArgAlaGluGlySerThrPr 611
 Db 2775 TGGAAACCCCAACCCCTCCATGGATGGAGAACTCTGCTGAGGGCAGAGGGATCTACGCC 2834
 Qy 611 oAlaGlyGlyLeuSerGlyGlyGlyPheGlnProSerGlyLeuAlaPheAlaSe 631
 Db 2835 AGCAGGTGGAGCTTGTTCAGGGGTGGGGCTTTTCAGGCCCTCTGGCTTGGCTTGTTC 2894
 Qy 631 rHisVal 633
 Db 2895 ACACGTG 2901

RESULT 8
 AX549295
 LOCUS AX549295 3119 bp DNA linear PAT 26-NOV-2002
 DEFINITION Sequence 580 from Patent WO02061087.
 ACCESSION AX549295
 VERSION AX549295.1 GI:25813958
 KEYWORDS
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1
 Burmer,G.C., Roush,C.L. and Brown,J.P.
 Antigenic peptides, such as for G protein-coupled receptors
 (GPCRs), antibodies thereto, and systems for identifying such
 antigenic peptides
 Patent: WO 02061087-A 580 08-AUG-2002;

FEATURES
 source Lifespan Biosciences, Inc. (US)
 Location/Qualifiers
 1..3119
 /organism="Homo sapiens"
 /mol_type="unassigned DNA"
 /db_xref="taxon:9606"
 ORIGIN
 Alignment Scores: 9.96e-208 Length: 3119
 Pred. No.: 3045.50 Matches: 602
 Score: 92.61% Conservative: 12
 Percent Similarity: 90.80% Mismatches: 19
 Best Local Similarity: 92.09% Indels: 32
 Query Match: 6 Gaps: 4
 DB: 4
 US-10-664-667-5 (1-633) x AX549295 (1-3119)
 Qy 1 AsnThrThrHisTyr----- 5
 Db 802 AACTCCACACACTATCTCTGAATGGTGCCATCCAGAGTTTCCAGATCTCAAG 861
 Qy 6 ---ArgGluSerTrpTyrAla-----CysArgTyrArg 15
 Db 862 GCACCACACGCTGGAGATCTGACCTGACCCGCGCAGGCATCCGGCTGCTCCCATCG 921
 Qy 16 -----SerGlyIleProGlySer-ThrHisAlaSerValGluArgSerGln---G1 31
 Db 922 GGATGTCCCAACAGCTGCCCAGGCTCCGAGTCTCGGAACCTGTCTCACAATCAAAATTGAGG 981
 Qy 31 YLeuSerLeuProAlaHisProAlaSerLeuAlaLeuAlaAlaSerAenThrThrAl 51
 Db 982 AGTGCCCGCCCTGCACAGGTGTGCAGAAATGGAGGAAA-TCGGCCTCCACACACCGC 1040
 Qy 51 aSerGlyLeuGlu***AspThrPheSerGlnLeuSerSerLeuGlnAlaLeuAspLe 71
 Db 1041 ATCTGGAAATTGGAGCT-GACACCTTCAGCAGCTGAGCTCCCTGCAAGCCCTGGATCT 1099
 Qy 71 uSerTrpAsnAlaIleArgSerIleHisProGluAlaPheSerThrLeuHisSerLeuVa 91
 Db 1100 TAGCTGGAAAGCCATCCGGTCCATCCACCTGAGGCCTTCTCCACCTGCACCTCCCTGGT 1159
 Qy 91 lLysLeuAspLeuThrAspAenGlnLeuThrThrLeuProLeuAlaGlyLeuGlyGlyLe 111
 Db 1160 CAAGCTGGACCTGCACAGAACACAGCTGACCACTGCCCTGGCTGACTTGGGGGCTT 1219
 Qy 111 uMetHisLeuLeuLysGlyAsnLeuAlaLeuSerGlnAlaPheSerLysAspSerPh 131
 Db 1220 GATGCATCTGAAGCTCAAAGGGAACCTTGCTCTCTCCAGGCCTTCTCCAAGGACAGTTT 1279
 Qy 131 eProLysLeuArgIleLeuGluValProTyrAlaTyrGlnCysCysProTyrGlyMetCy 151
 Db 1280 CCCAAACTGAGGATCTGGAGGTGCTTATGCTTACAGTGTCTCCCTATGGATGTG 1339
 Qy 151 sAlaSerPhePheLysAlaSerGlyGlnTrpGluAlaGluAspLeuHisLeuAspAspGl 171
 Db 1340 TGCCAGCTTCTTCAAGGCCTCTGGGAGTGGAGGCTGAAGACCTTCACCTTGATGATGA 1399
 Qy 171 uGluSerSerLysArgProLeuGlyLeuLeuAlaArgGlnAlaGluAenHisTyrAspGl 191
 Db 1400 GGAGTCTTCAAAAAGGCCCTGGGCTTCTTGGCAGACAAGCAGAGAACCACTATGACCA 1459
 Qy 191 nAspLeuAspGluLeuGlnLeuGluMetGluAspSerLysProHisProSerValGlnCy 211
 Db 1460 GGCCTGGATGAGCTCCAGCTGGAGATGGAGGACTCAAGGCCACACCCAGTGTCCAGTG 1519
 Qy 211 sSerProThrProGlyProPheLysProCysGluTyrLeuPheGluSerTrpGlyIleAr 231
 Db 1520 TAGCCCTACTCCAGGCCCTTCAAGCCCTGTGAGTACTCTTTGAAAGCTGGGGCATCG 1579
 Qy 231 gLeuAlaValTrpAlaIleValLeuLeuSerValLeuCysAsnGlyLeuValLeuLeuTh 251
 Db 1580 CCTGGCCGTGTGGGCCATCGTGTGCTCTCCGTGCTCTGCAATGAGAGTGTGTGCTGAC 1639


```
Db 1983 GCTTCGACGAGGGTCTAGCTGCTGCACTGGCAGGGCTGGCCGCGCACTGCCCT 2042
Qy 371 uAlaSerValGlyGluTyrGlyAlaSerProLeuCysLeuProTyrAlaProGluGl 391
Db 2043 GGCCTCAGTGGAGAAATACGGGGCTCCCACTCTGCTGCGCTACGCGCACCTGAGGG 2102
Qy 391 yGlnProAlaAlaLeuGlyPheThrValAlaLeuValMetMetAenSerPheCysPheLe 411
Db 2103 TCAGCCAGACGCCCTGGGCTTACCGTGCGCTTGTGTATGATGAACCTCTTCTGTTCCT 2162
Qy 411 uValValAlaGlyAlaTyrIleLysLeuTyrCysAspLeuProArgGlyAspPheGluAl 431
Db 2163 GGTCTGGCGGTGCCTACATCAACTGTACTGTGACCTGCGCGGGGAGCTTTGAGGC 2222
Qy 431 aValTrrpAspCysAlaMetValArgHisValAlaTrrpLeuIlePheAlaAspGlyLeuLe 451
Db 2223 CGTGTGGGACTGCGCCATGTGTGAGACAGTGGCTTGCCTCATCTTCGAGACGGGCTCCT 2282
Qy 451 uTyrCysProValAlaPheLeuSerPheAlaSerMetLeuGlyLeuPheProValThrPr 471
Db 2283 CTACTGTCTCCGCTGCTTCCCTCAGCTTGCCTTCCATGCTGGGCTTCTTCCCTGTACGCC 2342
Qy 471 oGluAlaValLysSerValLeuLeuValValLeuProLeuProAlaCysLeuAenProLe 491
Db 2343 CGAGCGGCTCAAGTCTGCTGCTGTGTGTGCTGCTGCCCTGCTGCTGCCTCAACCCACT 2402
Qy 491 uLeuTyrLeuLeuPheAenProHisPheArgAspAspLeuArgArgLeuArgProAla 511
Db 2403 GCTGTACCTGCTTCAACCCCACTTCCGGGATGACCTTCCGGGGCTTCCGGCCCGCGC 2462
Qy 511 aGlyAsnSerGlyProLeuAlaTyrAlaAlaAlaGlyGluLeuGluLysSerCysAs 531
Db 2463 AGGGGACTCAGGGGCCCTTAGCTTATGCTGCGCGCGGGAGCTGGAGAGACTCCTGTGA 2522
Qy 531 pSerThrGlnAlaLeuValAlaPheSerAspValAspLeuIleLeuGluAlaSerGluAl 551
Db 2523 TTCTACCCAGCCCTGGTAGCTTCTGTATGTGTGATCTCATTTGAGAGCTTCTGAGC 2582
Qy 551 aGlyArgProGlyLeuGluThrTyrGlyPheProSerValThrLeuLeuSerCysGl 571
Db 2583 TGGCGGGCCCCCTGGGCTGGAGACTATGGCTTCCCTCAGTGACCTCATCTCTGTCA 2642
Qy 571 nGlnProGlyAlaProArgLeuGluGlySerHisCysValGluProGluGlyAenHisPh 591
Db 2643 GCAGCCAGGGCCCCAGGCTGGAGGGCAGCATTTGTGTAGGCCAGAGGGGAACCACTT 2702
Qy 591 eGlyAsnProGlnProSerMetAspGlyGluLeuLeuLeuArgAlaGluGlySerThrPr 611
Db 2703 TGGGAACCCCAACCCCTCCATGGATGGAGAACTGCTGCTGAGGGCAGAGGGATCTACGCC 2762
Qy 611 oAlaGlyGlyGlyLeuSerGlyGlyGlyPheGlnProSerGlyLeuAlaPheAlaSe 631
Db 2763 AGCAGGTGGAGGCTTGTTCAGGGGGTGGCGGCTTTTCAGCCCTCTGGCTTGGCTTGTTC 2822
Qy 631 rHisVal 633
Db 2823 ACACGT 2829
```

RESULT 11

AX451929 LOCUS AX451929 3429 bp DNA linear PAT 03-JUL-2002
DEFINITION Sequence 26 from Patent WO0226825.

ACCESSION AX451929

VERSION AX451929.1 GI:21698752

SOURCE Homo sapiens (human)

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE

1 Baughn, M.R., Gaul, R.C., Wallia, N.K., Gandhi, A.R., Hafalia, A.J.,
Ramkumar, J., Tribouley, C.M., Thornton, M., Kallick, D.A., Yao, M.G.,

Elliott, V.S., Burford, N., Khan, F.A., Yue, H., Lu, Y., Arvizu, C.,
Roopa, R., Nguyen, D.B., Lee, E.A., Lu, D.A., Ison, C.H., Walsh, R.T. and
Policky, J.L.
G-protein coupled receptors
Patent: WO 0226825-A 26 04-APR-2002;
Incyte Genomics, Inc. (US)
FEATURES
source Location/Qualifiers
1..3429
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
/note="Incyte ID No: 7479890CB1"

ORIGIN

Alignment Scores:
Pred. No.: 111e-207 Length: 3429
Score: 3045.50 Matches: 602
Percent Similarity: 92.61% Conservative: 12
Best Local Similarity: 90.80% Mismatches: 19
Query Match: 92.09% Indels: 32
DB: 6 Gaps: 4

US-10-664-667-5 (1-633) x AX451929 (1-3429)

```
Qy 1 AsnThrThrHisTyr----- 5
Db 860 AACTCCACACACATATCTGTAATGGTGCCATGGACATCCAGGAGTTTCCAGGTTCTCAAG 919
Qy 6 ---ArgGluSerTrpTyrAla----- CysArgTyrArg 15
Db 920 GCACCACAGCTGGAGATCTGACCTGACCCGCGCAGGCATCCGGCTGCTCCCATCGG 979
Qy 16 -----SerGlyIleProGlySer-ThrHisAlaSerValGluArgSerGln---G1 31
Db 980 GGATGTGCTAACACAGCTGCCCAGGCTCCGAGTCTCGGAACCTGTCTCACAATCAAAATTGAGG 1039
Qy 31 yLeuSerLeuProAlaHisProAlaSerLeuAlaLeuAlaAlaSerAenThrAl 51
Db 1040 AGCTGCCAGCCTGCACAGGTGTCAGAAATTTGGAGGAAA--TCGGCCTTCCAAACACACCGC 1098
Qy 51 aSerGlyIleLeuGlu***AspThrPheSerGlnLeuSerSerLeuGlnAlaLeuAspLe 71
Db 1099 ATCTGGGAATTTGAGCT--GACACCTTTCAGCCAGCTGAGCTCCCTGCAAGCCCTGGATCT 1157
Qy 71 uSerTrpAsnAlaIleArgSerIleHisProGluAlaPheSerThrLeuHisSerLeuVa 91
Db 1158 TAGCTGGAACGCCATCCGGTCCATCCACCCGAGGCCCTTCTCCACCTGCACTCCCTGGT 1217
Qy 91 lLysLeuAspLeuThrAspAenGlnLeuThrThrLeuProLeuAlaGlyLeuGlyLe 111
Db 1218 CAAGCTGGACCTACACAGAACCCAGCTGACACACACTGCCCTGGCTGGACTTGGGGGCTT 1277
Qy 111 uMetHisLeuLysLeuLysGlyAenLeuAlaLeuSerGlnAlaPheSerLysAspSerPh 131
Db 1278 GATGATCTCAAGCTCAAAAGGAAACCTTGTCTCTCCAGGCCCTTCTCCAAGGACAGTTT 1337
Qy 131 eProLysLeuArgIleLeuGluValProTyrAlaTyrGlnCysCysProTyrGlyMetCy 151
Db 1338 CCCAAAACCTGAGGATCCTGGAGGTGCTTATGCTTACAGTGTGCTGCTTATGGGATGTG 1397
Qy 151 sAlaSerPhePheLysAlaSerGlyGlnTrpGluAlaGluAspLeuHisLeuAspAspGl 171
Db 1398 TGCAGGCTTCTTCAAGGCTTCTGGGCAGTGGAGGCTGAAGACTTCAACCTTGNATGATGA 1457
Qy 171 uGluSerSerLysArgProLeuGlyLeuLeuAlaArgGlnAlaGluAenHisTyrAspGl 191
Db 1458 GGAGTCTTCAAAAAGGCCCCCTGGGCTCTCTTCAGAGCAAGCAGAGAACCACTATGACCA 1517
Qy 191 nAspLeuAspGluLeuGlnLeuGluMetGluAspSerLysProHisProSerValGlnCy 211
Db 1518 GGACCTGGATGAGCTCAGCTGAGATGGAGGACTCAAGAGCCACACCCCACTGATGCCAGTG 1577
Qy 211 sSerProThrProGlyProPheLysProCysGluTyrLeuPheGluSerTrpGlyIleAr 231
```


Db 1578 TAGCCCTACTCCAGGCCCTTCAAGCCCTGTGAGTACCTCTTTGAAAGCTGGGGCATCCG 1637
 Qy 231 gLeuAlaValTAlaIleValLeuSerValLeuCysAnGlyLeuValLeuTh 251
 Db 1638 CTGGCGGTGGGCGCATCGTGTCTCCGTCTCTGCATGACGTGCTGCTGAC 1697
 Qy 251 rValPheAlaGlyGlyProAlaProLeuProValValysPheValValGlyAlaIleAl 271
 Db 1698 CGTGTTCGTGGCGGCGCTGCCCTGCCCGCGGTCAAGTTTGTGTAGGTGCGATTGC 1757
 Qy 271 aGlyAlaSerThrLeuThrGlyLysSerCysGlyLeuLeuAlaSerValAspAlaLeuTh 291
 Db 1758 AGGCCCCACACCTTGACTGGCATTTCTGTGCGCTTCTAGCCTCAGTCGATGCCCTGAC 1817
 Qy 291 rPheGlyGlnPheSerGluTyrGlyAlaArgTrpGluThrGlyLysCysArgAlaTh 311
 Db 1818 CTTTGGTCAGTTCTCTGAGTACGGAGCCGCTGGAGACGGGCTAGGCTGCCGGGCCAC 1877
 Qy 311 rGlyPheLeuAlaValLeuGlySerGluAlaSerValLeuLeuLeuThrLeuAlaAlaVa 331
 Db 1878 TGGCTTCCTGGCAGTACTTGGGTTCGGAGGCATCGTGTCTGCTCACTCTGCGCGCAGT 1937
 Qy 331 lGlnCysSerValSerValSerCysValArgAlaTyrGlyLysSerProSerLeuGlySe 351
 Db 1938 CGAGTGCAGCGTCTCCGTCTCTGTGTCGGGCGCTATGGAAAGTCCCTCCCTCGGCGAG 1997
 Qy 351 rValArgAlaGlyValLeuGlyCysLeuAlaLeuAlaGlyLeuAlaAlaLeuProle 371
 Db 1998 CBTTCGAGCGGGTCTTAGGCTGCTGGCATTCGGAGGGCTGGCGCGGCACTCCCTCC 2057
 Qy 371 uAlaSerValGlyGluTyrGlyAlaSerProLeuCysLeuProTyrAlaProProGluGl 391
 Db 2058 GSCTTCAGTGGGAGTAACGGGGCTCCCACTCTGCTGCTGCTACGGCGCACCTGAGGG 2117
 Qy 391 yGlnProAlaAlaLeuGlyPheThrValAlaLeuValMetMetAnSerPheCysPheLe 411
 Db 2118 TCAGCAGCAGCGCTGGGCTTCACCGTGGCGCTTGTGTGATGATGAACCTCTCTGTTCCT 2177
 Qy 411 uValValAlaGlyAlaTyrIleLeuLeuTyrCysAspLeuProArgGlyAspPheGluAl 431
 Db 2178 GTCGTGGCGGTGCTACATACTGATGATGATGATGATGATGATGATGATGATGATGATG 2237
 Qy 431 aValTrpAspCysAlaMetValArgHisValAlaTrpLeuIlePheAlaAspGlyLeuLe 451
 Db 2238 CGTGTGGGACTGGCGCATGGTGAGCAGCGTGGCTGGCTCATCTTCGACAGCGGCTCT 2297
 Qy 451 uTyrCysProValAlaPheLeuSerPheAlaSerMetLeuGlyLeuPheProValThrPr 471
 Db 2298 CTACTGTCCGTGGGCTTCTCAGCTTTGCTCCATGCTGCTGCTGCTGCTGCTGCTGCT 2357
 Qy 471 oGluAlaValLysSerValLeuValLeuValLeuProLeuProAlaCysLeuAsnProle 491
 Db 2358 CGAGCGCGTCAAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2417
 Qy 491 uLeuTyrLeuLeuPheAnProHisPheArgAspLeuArgArgLeuArgProArgAl 511
 Db 2418 GCTGTACTGCTCTTCAACCCCTTTCGGGATGACCTTCGGCGGCTTCGGCGGCGGCG 2477
 Qy 511 aGlyAspSerGlyProLeuAlaTyrAlaAlaAlaGlyGluLeuGluLysSerSerCysAs 531
 Db 2478 AGGGGACTCAGGGGCCCTAGCCTATGCTGCGCGCGGCGGAGTGGAGAGAGCTCTGTGA 2537
 Qy 531 pSerThrGlnAlaLeuValAlaPheSerAspValAspLeuLeuLeuAlaSerGluAl 551
 Db 2538 TTCTACCCAGGCGCTGGTAGCCTTCTGTATGATGATGATGATGATGATGATGATGATG 2597
 Qy 551 aGlyArgProProGlyLeuGluThrTyrGlyPheProSerValThrLeuIleSerCysG 571
 Db 2598 TGGCGGCGGCGCTGGGCTGGAGACTATGGCTTCCCTCCAGTGCCTCATCTCTCTGCA 2657
 Qy 571 nGlnProGlyAlaProArgLeuGluGlySerHisCysValGluProGluGlyAsnHisPh 591

Db 2658 GCAGCCAGGGGCCCCCAGGCTGGAGGGCAGCCATTGTGTAGAGCCAGAGGGGAACACACTT 2717
 Qy 591 eGlyAsnProGlnProSerMetAspGlyGluLeuLeuArgAlaGluGlySerThrPr 611
 Db 2718 TGGGNAACCCCAACCTTCATGGATGGAGAACTGCTGCTGAGGGGAGAGGATCTACGCC 2777
 Qy 611 oAlaGlyGlyGlyLeuSerGlyGlyGlyGlyPheGlnProSerGlyLeuAlaPheAlaSe 631
 Db 2778 AGCAGGTGGAGCTTGTTCAGGGGGTGGCGGCTTTTCAGGCCCTCTGCTTGGCCTTTGCTTC 2837
 Qy 631 rHisVal 633
 Db 2838 ACACGTG 2844
 RESULT 12
 LOCUS AX301831 3492 bp DNA linear PAT 30-NOV-2001
 DEFINITION Sequence 10 from Patent WO0185768.
 ACCESSION AX301831
 VERSION AX301831.1 GI:17382891
 KEYWORDS Homo sapiens (human)
 SOURCE
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1
 AUTHORS Gu, W.
 TITLE Novel g-protein coupled receptors and uses therefor
 JOURNAL Patent: WO 0185768-A 10 15-NOV-2001;
 Millennium Pharmaceuticals, Inc. (US)
 FEATURES
 source
 1. 3492
 /organism="Homo sapiens"
 /mol_type="unassigned DNA"
 /db_xref="taxon:9606"
 104..3007
 /note="unnamed protein product"
 /codon_start=1
 /protein_id="CAD13081.1"
 /db_xref="GI:17382892"
 /translation="MPSPPGLRLWLCAALCASRRAGAPQPGPPTACAPACHQED
 GIMUADCSSELGSAVFGDLDPLTAYLDLSMNNLTQLQGLFHLRLEELRLSGNHL
 SHIPQAFVSLKILMNNQGGIPAEALWELPSQLSLRDLNANLISLVPSRFEFEG
 LSSLRHLWLDNALTEIPVRLNLPALQAMTLALNRISHIPDYAFONLTSLVVLHLH
 NNRQHLGTHSFEGLNLETLDLNVLKLOEPPVAIRTLGRLOELGFHNNTKAIPEKA
 PMGNPLQTHFYDNPITQVGRSAFYQLKLTLSLNGAMDIQFPDCLKGTTLSLILT
 LTRAGIRLLPSGMCQQLPRURVLESHNQIEELPSLHRCQKLEIGLQHNWEIGAD
 TFSQSLQLADLSWNAIRSIHPEAFSTLSLVKCLDLTDNQLTTLPLAGGLMLHLK
 KGNIALSQAFSKDSFKRLRIIEVYAYCCPYGMCASFPFKASGQWEABDLHLDDESS
 KRPIGLLARQAEHVDLDLELEMEDESKPHSVQCSPTGPKPKCYLPSMWGIRL
 AVMAIVLLVLCNGLVLLTVPAGGPALPPKPVVGAAGANTLITGSCGLASVDAL
 TFGQFSEYGRWETGLCRATGFLAVLGSEASVLLTLAAVQCSVSCVRAYGKSPS
 LGSVRAGVLCGLALAGLAALPLASVEYGASPLCLPYAPEGQPALGFTVALVMMN
 SFCEVLVAGAYIKLYCDLPRGDFEAVDCAMVRHAWLIFADGLLYCPVAFSLFMSML
 LPTPTPEAVKSVLLVLLYLLPACLNPLLYLFPNPHFDRLRLRPRAGDSGLAYAAA
 GELFKSSCDSTQALVAFSDVLDLLEASEAGRPGLETYGPSPVTLISQQQGPARGLEG
 SHCPEPGNHFNGPQPSMDGELLRLRAEGSTPAGGGLSGGGGFGQSGGLAFASHV"
 ORIGIN
 Alignment Scores:
 Pred. No.: 1,14e-207 Length: 3492
 Score: 3045.50 Matches: 602
 Percent Similarity: 92.61% Conservativeness: 12
 Best Local Similarity: 90.80% Mismatches: 19
 Query Match: 92.09% Indels: 32
 DB: 6 Gaps: 4
 US-10-664-667-5 (1-633) x AX301831 (1-3492)
 Qy 1 AsnThrThrHisTyr-----
 Db 1020 AACTCCACACTATCTCTGAATGGTGCATGGACATCCAGGAGTTTCCAGATCTCAAAG 1079

Qy 6 ---ArgGluSerTrpTyrAla-----CysArgTyrArg 15
 Db 1080 GCACACAGCCTGGAGATCTGACCCCTGACCCGCGCAGGAGATCGGCTGCTCCATCGG 1139
 Qy 16 -----SerGlyIleProGlySer-ThrHisAlaSerValGluArgSerGln---G1 31
 Db 1140 GGATGTGCCAACAGCTGCCAGGCTCCGAGTCTCGGAACACTCTCACAAATCAAAATTGAGG 1199
 Qy 31 YLeuSerLeuProAlaHisProAlaSerLeuAlaAlaLeuAlaAlaSerHisThrAl 51
 Db 1200 AGCTGCCAGCCTGCACAGGTGTGAGAAATGSGAGAAA-TCGGCCCTCCACACAAACCGC 1258
 Qy 51 aSerGlyIleLeuGlu***AspThrPheSerGlnLeuSerLeuGlnAlaLeuAspLe 71
 Db 1259 ATCTGGGAATTTGAGCT-GACACCTTCAGCCAGCTGAGCTCCCTGCNAGCCCTGGATCT 1317
 Qy 71 uSerTrpAsnAlaIleArgSerIleHisProGluAlaPheSerThrLeuHisSerLeuVa 91
 Db 1318 TAGCTGGAACGCCATCGGTCCATCCACCTGAGGCTTCTCCACCCCTGCACCTCCCTGGT 1377
 Qy 91 lLysLeuAspLeuThrAspAsnGlnLeuThrLeuProLeuAlaGlyLeuGlyLe 111
 Db 1378 CAAGCTGACCTGACAGACACACAGCTGACACACTGCCCTGGCTGGACTTGGGGGCTT 1437
 Qy 111 uMetHisLeuIleLeuIleGlyAsnLeuAlaLeuSerGlnAlaPheSerIleAspSerPh 131
 Db 1438 GATGCACTGGAAGCTCAAAAGGAACCTTGCTCTCTCCAGGCTTCTCCAGGACAGTTT 1497
 Qy 131 eProLysLeuArgIleLeuGluValProTyrAlaTyrGlnCysCysProTyrGlyMetCy 151
 Db 1498 CCCAAAACCTGAGGATCCTGGAGGTGCTTATGCTTACCAGTGCTGCTATGGATGTG 1557
 Qy 151 sAlaSerPhePheIleAlaSerGlyGlnTrpGluAlaGluAspLeuHisLeuAspG1 171
 Db 1558 TGCCAGCTTCTCAAGGCTCTGGGCAGTGGGAGGCTGAAGACCTTCACTTGTATGTA 1617
 Qy 171 uGluSerSerIleArgProLeuGlyLeuAlaAlaArgGlnAlaGluAsnHisTyrAspG1 191
 Db 1618 GGAGTCTTTCAAAAGGCCCTTGGGCTCTTGGCCAGACAGCAGAACCACTATGACCA 1677
 Qy 191 nAspLeuAspGluLeuGlnLeuGluMetGluAspSerIleProHisProSerValGlnCy 211
 Db 1678 GGACCTGATGAGCTCCAGCTGGAGATGGAGGACTCAAAAGCCACACCCAGTGTCCAGTG 1737
 Qy 211 sSerProThrProGlyProPheIlePheIlePheIlePheIlePheIlePheIleAr 231
 Db 1738 TAGCCCTACTCCAGGCCCTTCAAGCCCTGTGAGTACCTCTTTGAAAGCTGGGGCATCCG 1797
 Qy 231 gLeuAlaValTrpAlaIleValLeuLeuSerValLeuCysAsnGlyLeuValLeuLeuTh 251
 Db 1798 CTTGGCCGTGTGGGCCATCGTGTGCTCTCCGTGCTCTGCAATGGACTGGTGTGCTGAC 1857
 Qy 251 rValPheAlaGlyGlyProAlaProLeuProProValIlePheValValGlyAlaIleAl 271
 Db 1858 CGTGTGCTGGCGGCCCTGCCCTGCCCTGCCCTGCCCTGCCCTGCCCTGCCCTGCCCTGC 1917
 Qy 271 aGlyAlaAsnThrLeuThrGlyIleSerCysGlyLeuLeuAlaSerValAlaLeuTh 291
 Db 1918 AGGGCCCAACACCTTGATGGCAATTTCTGTGGCTTCTTAGCCCTCAGTCCAGTGGCCCTGAC 1977
 Qy 291 rPheGlyGlnPheSerGluTyrGlyAlaArgTrpGluThrGlyLeuGlyCysArgAlaTh 311
 Db 1978 CTTTGGTTCAGTCTCTGAGTACGAGCCCGCTGGAGACGGGGCTAGGCTGCCGGGCCAC 2037
 Qy 311 rGlyPheLeuAlaValLeuGlySerGluAlaSerValLeuLeuLeuThrLeuAlaVa 331
 Db 2038 TGGCTCTCCGTCAGTACTTGGGTCCGAGGATCGTGTGCTGTCTCACTCTGGCCGAGT 2097
 Qy 331 lGlnCysSerValSerValSerCysValArgAlaTyrGlyIleSerProSerLeuGlySe 351
 Db 2098 GCAGTGCAGCGTCTCCGCTCTCTGTGTCGGGCTATGGGAAGTCCCTCCCTGGGCGAG 2157
 Qy 351 rValArgAlaGlyValLeuGlyCysLeuAlaAlaGlyLeuAlaAlaLeuProLe 371

Db 2158 CGTTCGAGCAGGGGTCTAGGCTGCTGGCAGCTGGCAGGGCTGGCCCGCAGCTGCCCT 2217
 Qy 371 uAlaSerValGlyGlyTyroGlyAlaSerProLeuCysLeuProTyrAlaProProGluG1 391
 Db 2218 GGCTTCAGTGGGAGATACCGGGGCTCCCCACCTGCTGCCCTPACCGGCACCTGAGGG 2277
 Qy 391 YGlnProAlaAlaLeuGlyPheThrValAlaLeuValMetMetMetMetMetMetMet 411
 Db 2278 TCAGCCAGCAGCCTGGGCTTACCGTGGCCCTGGTGGTGGTGGTGGTGGTGGTGGTGG 2337
 Qy 411 uValValAlaGlyAlaTyroIleLysLeuTyroCysAspLeuProArgGlyAspPheGluAl 431
 Db 2338 GGTGCTGGCGGTGCCTTACATCAAACTGTACTGTGACCTGCCCGGGGCGACTTGGGC 2397
 Qy 431 aValTrpAspCysAlaMetValArgHisValAlaTrpLeuIlePheAlaSerGlyLeuLe 451
 Db 2398 CGTGTGGGACTGGGCCATGTTGAGGCAGCTGGGCTGGCTCATCTTCGACAGCGGGCTCT 2457
 Qy 451 uTyrCysProValAlaPheLeuSerPheAlaSerMetMetMetMetMetMetMetMet 471
 Db 2458 CTACTGTCCGCTGGCTTCTCAGCTTCCGCTTCCATGCTGGGCTCTTCCCTGTGCAGCC 2517
 Qy 471 oGluAlaValIleSerValLeuLeuValValLeuProLeuProAlaCysLeuAsnProLe 491
 Db 2518 CGAGCGGTCAAGTCTGCTGCTGCTGGTGTGCTGCCCTGCTGCCCTGCCTCAACCCACT 2577
 Qy 491 uLeuTyrLeuLeuPheAsnProHisPheArgAspAspLeuArgArgLeuArgProArgAl 511
 Db 2578 GCTGTACTCTCTTCAACCCCACTTCCGGGATGACCTTTCGGCGGCTTCGGCCCCCGGC 2637
 Qy 511 aGlyAspSerGlyProLeuAlaTyrAlaAlaAlaGlyLeuGlyLysSerSerCysAs 531
 Db 2638 AGGGACTCAGGGCCCCCTAGCTTATGCTGGCGCGGAGCTGGAGAGAGCTCTCTGTGA 2697
 Qy 531 pSerThrGlnAlaLeuValAlaPheSerAspValAspLeuIleLeuGluAlaSerGluAl 551
 Db 2698 TTCTACCCAGCCCTGGTGGCTTCTCTGATGTGGATCTCATCTTGAAGCTTCTGAAGC 2757
 Qy 551 aGlyArgProGlyLeuGluThrTyrGlyPheProSerValThrLeuIleSerCysG1 571
 Db 2758 TGGCGCGCCCTGGGCTGGAGACCTATGGCTTCCCTCAGTGACCTCATCTCTGTCA 2817
 Qy 571 nGlnProGlyAlaProArgLeuGlySerHisCysValGluProGluGlyAsnHisPh 591
 Db 2818 GCAGCCAGGGCCCCCAGGCTGGAGGCGACCAATTGTGTAGAGCCAGAGGGGACCACTT 2877
 Qy 591 eGlyAsnProGlnProSerMetAspGlyGluLeuLeuLeuAlaGluGlySerThrPr 611
 Db 2878 TGGGAACCCCAACCCCTCCATGGATGGAGAACTGCTGCTGAGGGCAGAGGATCTAGGCC 2937
 Qy 611 oAlaGlyGlyLeuSerGlyGlyGlyPheGlnProSerSerGlyLeuAlaPheAlaSe 631
 Db 2938 AGCAGGTGGAGGCTTGTGAGGGGCTGGCGCTTTCAGCCCTCTGGCTTGGCTTGGCTTC 2997
 Qy 631 rHisVal 633
 Db 2998 ACAGTG 3004
 RESULT 13
 LOCUS CO870634 3273 bp DNA linear PAT 13-SEP-2004
 DEFINITION Sequence 43 from Patent WO2004073657.
 ACCESSION CO870634
 VERSION CO870634.1 GI:52000145
 KEYWORDS
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1
 Aziz,N., Gish,K.C., Wilson,K.E. and Zlotnik,A.
 METHODS of diagnosis of cancer, composition and methods of

screening for modulators of cancer
Patent: WO 2004073657-A 43 02-SEP-2004;
PROTEIN DESIGN LABS, INC. (US)
FEATURES
source
1. 3273
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

ORIGIN

Alignment Scores:
Pred. No.: 2,048-207 Length: 3273
Score: 3041.50 Matches: 601
Percent Similarity: 92.46% Conservative: 12
Best Local Similarity: 90.65% Mismatches: 20
Query Match: 91.97% Indels: 32
Dbs: 6 Gaps: 4

US-10-664-667-5 (1-633) x CQ870634 (1-3273)

Qy 1 AsnThrThrHisTyr----- 5
Db 804 AACTCCACACACTATCTCTGAATGTGCCATGCATCCAGGAGTTTCCAGATCTCAAAG 863
Qy 6 ---ArgGluSerTrpTyrAla-----CysArgTyrArg 15
Db 864 GCACACACAGCTGGAGATCTCTGACCTCGACCCGCGCAGGATCCGGCTGCTCCCATCGG 923
Qy 16 -----SerGlyIleProGlySer-ThrHisAlaSerValGluArgSerGln--G1 31
Db 924 GGATGTGCCAACAGCTGCCAGGCTCCGAGTCTCGAACTGCTCACAATCAAAATTGAGG 983
Qy 31 yLeuSerLeuProAlaHisProAlaSerLeuAlaAlaLeuAlaAlaSerAsnThrAl 51
Db 984 AGCTGCCAGCTGCACAGGTCTCAGAAATTGGAGAAA-TCGGCTCCACACACACCGC 1042
Qy 51 aSerGlyIleLeuGlu***AspThrPheSerGlnLeuSerSerLeuGlnAlaLeuAspLe 71
Db 1043 ATCTGGGAAATTGGAGCT-GACACCTTCAGCCAGCTGAGTCCCTGCAAGCCCTGGATCT 1101
Qy 71 uSerTrpAlaIleArgSerIleHisProGluAlaPheSerThrLeuHisSerLeuVa 91
Db 1102 TAGCTGGAAAGCCATCCCGGTTCATCCACCCGAGGCTTCTCCACCTGCACCTCCCTGGT 1161
Qy 91 llysLeuAspLeuThrAspAsnGlnLeuThrThrLeuProLeuAlaGlyLeuGlyLe 111
Db 1162 CAAGCTGACCTGCACAGACACACAGCTGACACACTGCCCTGGCTGAGCTTGGGGCTT 1221
Qy 111 uMetHisLeuIleLeuSerGlyAsnLeuAlaLeuSerGlnAlaPheSerLysAspSerPh 131
Db 1222 GATGCATCTGAAGCTCAAAGGGAACCTTGCTCTCCAGGCTTCTCCAGGACAGTTT 1281
Qy 131 eProlysisLeuArgIleLeuGluValProTyrAlaTyrGlnCysCysProTyrGlyMetCy 151
Db 1282 CCCAAAATGAGGATCTCTGGAGGTGCTTATGCTTACCAGTGCTGCTCCCTATGGGATGTG 1341
Qy 151 sAlaSerPhePheTyrAlaSerGlyGlnTrpGluAlaGluAspLeuHisLeuAspAspG1 171
Db 1342 TGCCAGCTTCTCAAGGCTCTGGGAGCTGGAGGCTGAGACCTTACCTTGATGAIGA 1401
Qy 171 uGluSerSerLysArgProLeuGlyLeuAlaAlaArgGlnAlaGluAsnHisTyrAspG1 191
Db 1402 GGAGTCTTCAAAGGCCCTGGGCTCTCTGGCAGACAGCAGAGAACCACTATGACCA 1461
Qy 191 nasLeuAspGluLeuGlnLeuGluMetGluAspSerLysProHisProSerValGlnCy 211
Db 1462 GGACCTGATGAGCTCCAGCTGGAGATGGAGGACTCAAAGGCCACACCCAGTGTCCAGTG 1521
Qy 211 sSerProThrProGlyProPheLysProCysGlnTyrLeuPheGluSerTrpGlyIleAr 231
Db 1522 TAGCCCTACTCCAGGCCCTTCAAGCCCTGTGAGTACCTCTTTGAAAGCTGGGGATCCG 1581
Qy 231 gLeuAlaValTrpAlaIleValLeuLeuSerValLeuCysAsnGlyLeuValLeuLeuTh 251

Db 1582 CTGGCCCTGTGGGCCATCGTGTGCTCTCGCTGCTCTCGCAATGACGTGGTCTGCTGAC 1641
Qy 251 rValPheAlaGlyGlyProAlaProLeuProProValLysPheValValGlyAlaIleAl 271
Db 1642 CGTGTTCGTGGGGGCTGTCCCTGCCCGGTCAAGTTTGTGGTAGGTGCCATTCG 1701
Qy 271 aGlyAlaAsnThrLeuThrGlyIleSerCysGlyLeuLeuAlaSerValAspAlaLeuTh 291
Db 1702 AGGCGCAACACCTTGACTGGCATTTCTGTGGCCTTCTAGCCTCAGTCGATGCCCTGAC 1761
Qy 291 rPheGlyGlnPheSerGluTyrGlyAlaArgTrpGluThrGlyLeuGlyCysArgAlaTh 311
Db 1762 CTTTGGTCAGTTCTCTGAGTACGGAGCCGCTGGGAGACGGGGCTAGGCTGCGGGCCAC 1821
Qy 311 rGlyPheLeuAlaValLeuGlySerGluAlaSerValLeuLeuLeuThrLeuAlaAlaVa 331
Db 1822 TGGCTTCTCGCAGTACTTGGGTGGAGGCATCGGTGCTGCTGCTCATCTTGGCCGCGAGT 1881
Qy 331 lGlnCysSerValSerValSerCysValArgAlaTyrGlyLysSerProSerLeuGlySe 351
Db 1882 GCAGTGCAGCGTCTCCGTCTCTGTGTCGGGCTTATGGGAAGTCCCTCTCCCTGGGCGAG 1941
Qy 351 rValArgAlaGlyValLeuGlyCysLeuAlaLeuAlaGlyLeuAlaAlaLeuProLe 371
Db 1942 CGTTTCGAGCAGGGGTCTTAGGCTGCTGGCCTGCTGGCAGGCTGGCGCGCGCTCCCT 2001
Qy 371 uAlaSerValGlyGluTyrGlyAlaSerProLeuCysLeuLeuProTyrAlaProProGluG1 391
Db 2002 GGCTTCAGTGGGAAATACGGGGCTCTCCCACTCTGCTGCTGCTACGGCGCACCTGAGGG 2061
Qy 391 yGlnProAlaAlaLeuGlyPheThrValAlaLeuValMetMetAsnSerPheCysPheLe 411
Db 2062 TCAGCCAGCAGCTGGGCTTCCCGTGGCTTCCGTGATGATGAACCTCTTCTGTTCCT 2121
Qy 411 uValValAlaGlyAlaTyrIleLysLeuTyrCysAspLeuProArgGlyAspPheGluAl 431
Db 2122 GGTGTCGGCGGTGCTACATCAAACTGTACTGTGACCTGCGCGGGCGAGCTTTGAGGC 2181
Qy 431 aValTrpAspCysAlaMetValArgHisValAlaTrpLeuIlePheAlaAspGlyLeuLe 451
Db 2182 CGTGGGAGCTGCGCCATGGTGAGGACGCTGGCTGCTCATCTTCGAGACGGGCTCCT 2241
Qy 451 uTyrCysProValAlaPheLeuSerPheAlaSerMetLeuGlyLeuPheProValThrPr 471
Db 2242 CTACTGTCCTGGCTTCCCTCAGCTTTCCTCCATGCTGGGCTCTTCCCTGTGTCACGCC 2301
Qy 471 oGluAlaValLysSerValLeuLeuValValLeuProLeuProAlaCysLeuAsnProLe 491
Db 2302 CGAGCGCTCAAGTCTGTCTGCTGGTGTGCTGCTGCTGCTGCTGCTGCTCAACCCACT 2361
Qy 491 uLeuTyrLeuLeuPheAsnProHisPheArgAspLeuArgLeuArgLeuArgProArgAl 511
Db 2362 GCTGTACTGCTCTTCAACCCCTTCCGGATGACCTTCGGGGGCTTCGGCCCCGGCG 2421
Qy 511 aGlyAspSerGlyProLeuAlaTyrAlaAlaAlaGlyLysLeuLysSerSerCysAs 531
Db 2422 AGGGAGCTCAGGGCCCTTAGCTATGCTGCGCGGGGAGCTGGAGAGAGAGCTCTCTGTA 2481
Qy 531 pSerThrGlnAlaLeuValAlaPheSerAspValAspLeuIleLeuGluAlaSerGluAl 551
Db 2482 TTCTACCGCGCTGGTAGCTTCTCTGATGTGATCTCATTTCTGGAAGCTTCTGGAAGC 2541
Qy 551 aGlyArgProProGlyLeuGluThrTyrGlyPheProSerValThrLeuIleSerCysG1 571
Db 2542 TGGCGCGCCCTGGGCTGGAGACTTATGGCTTCCCTCAGTGACCTCTCTCTCTGTA 2601
Qy 571 nGlnProGlyAlaProArgLeuGluGlySerHisCysValGluProGluGlyAsnHisPh 591
Db 2602 GCAGCCAGGGCCCCCAGGCTGGAGGGCAGCCATTGTGTAGAGCCAGAGGGGACCACTT 2661
Qy 591 eGlyAsnProGlnProSerMetAspGlyGluLeuLeuLeuArgAlaGluGlySerThrPr 611

Db	2662	TGGGAACCCCAACCTCCATGGATGGAGAACTGCTGCTGAGGGGAGAGGATCTACGCC	2721
Qy	611	oAlaGlyGlyGlyLeuSerGlyGlyGlyGlyPheGlnProSerGlyLeuAlaPheAlaSe	631
Db	2722	AGCAGTGGAGGCTTGTGAGGGGGTGGCGCTTTCAGCCCTCTGGCTTGGCTTTC	2781
Qy	631	rHisVal	633
Db	2782	ACACGTG	2788
RESULT 14	BC047905	3286 bp mRNA linear PRI 22-APR-2003	
LOCUS	BC047905	Homo sapiens leucine-rich repeat-containing G protein-coupled	
DEFINITION	receptor 6, mRNA (CDNA clone IMAGE:5212508), partial cds.		
ACCESSION	BC047905		
VERSION	BC047905.1	GI:29126870	
KEYWORDS			
SOURCE	Homo sapiens (human)		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
AUTHORS	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.		
	1 (bases 1 to 3286)		
	Strausberg, R.D., Collins, F.S., Wagner, L.H., Grouse, L.H., Derge, J.G.,		
	Klausner, R.D., Collings, F.S., Wagner, L.H., Shenmen, C.M., Schuler, G.D.,		
	Altschul, S.F., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Bhat, N.K.,		
	Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Hsieh, F.,		
	Diatchenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L.,		
	Stapleton, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L.,		
	Scheetz, T.E., Brownstein, M.J., Usdin, T.B., Toshiyuki, S.,		
	Carninci, P., Prange, C., Raha, S.S., Loquellano, N.A., Peters, G.J.,		
	Abramson, R.D., Mullahy, S.J., Bosak, S.A., McEwan, P.J.,		
	McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S.,		
	Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hulyk, S.W.,		
	Villalón, K.C., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A.,		
	Fahey, J., Helton, E., Kettner, M., Madan, A., Rodriguez, S.,		
	Sanchez, A., Whiting, M., Madan, A., Young, A.C., Shevchenko, Y.,		
	Bouffard, G.G., Blakesley, R.W., Touchman, J.W., Green, E.D.,		
	Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M.,		
	Butterfield, Y.S., Krzywinski, M.I., Skalska, U., Smalusz, D.E.,		
	Schneer, A., Schein, J.E., Jones, S.J. and Marra, M.A.		
TITLE	Generation and initial analysis of more than 15,000 full-length		
	human and mouse cDNA sequences		
JOURNAL	Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)		
MEDLINE	22388257		
PUBMED	12477932		
REFERENCE	2 (bases 1 to 3286)		
AUTHORS	Strausberg, R.		
TITLE	Direct Submission		
JOURNAL	Submitted (03-MAR-2003) National Institutes of Health, Mammalian		
	Gene Collection (MGC), Cancer Genomics Office, National Cancer		
	Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,		
	USA		
REMARK	NIH-MGC Project URL: http://mgc.nci.nih.gov		
COMMENT	Contact: MGC help desk		
	Email: cgapbs-remail.nih.gov		
	Tissue Procurement: Life Technologies, Inc.		
	cDNA Library Preparation: Life Technologies, Inc.		
	cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)		
	DNA Sequencing by: Sequencing Group at the Stanford Human Genome		
	Center, Stanford University School of Medicine, Stanford, CA 94305		
	Web site: http://www-shgc.stanford.edu		
	Contact: (Dickson, Mark) mcd@paxil.stanford.edu		
	Dickson, M., Schmutz, J., Grimwood, J., Rodriguez, A., and Myers,		
	R. M.		
Clone distribution:	MGC clone distribution information can be found		
through the I.M.A.G.E. Consortium/LLNL at:	http://image.llnl.gov		
Series:	IRAK Plate: 107 Row: d Column: 11.		
Location/Qualifiers	1. 3286		
/organism="Homo sapiens"			
/mol_type="mRNA"			

Db	1145		CAAGCTGACCTGACAGACAAACCAAGCTGACACACTGCCCTGGCTGGACTTGGGGCTT	1204
Qy	111		uMetHisLeuLeuysGlyvAanLeuAlaLeuSerGlnAlaPheSerLyAspSerPh	131
Db	1205		GATGCATCTGAAGCTCAAGGGACCTTGCTCTCTCCAGGCCCTTCTCCAAGACAGTTT	1264
Qy	131		eProLyLeuArgIleLeuGluValProTyrAlaTyrglnCysCysProTyrglyMetCy	151
Db	1365		CCCCAAACTGAGGATCTCTGGAGGTGCCTTATGCCTACCAAGTGTCTCCCTATGGGATGTG	1324
Qy	151		eAlaSerPhePheLyAlaSerGlyGlnTrpGluAlaGluAspLeuHisLeuAspAspGl	171
Db	1325		TGCCAGCTTCTTCAAGGCCCTCTGGCAGTGGAGGCTCAAGACCTTCACTGTGATGATGA	1384
Qy	171		uGluSerSerLyAspArgProLeuGlyLeuLeuAlaArgGlnAlaGluAenHisTyAspGl	191
Db	1385		GGAGCTCTCAAAAAGGCCCTTGGGCCCTCTTGCCAGACAGCAGAACCACTATGACCA	1444
Qy	191		nAspLeuAspGluLeuGlnLeuGluMetGluAspSerLyAspHisProHisProSerValGlnCy	211
Db	1445		GGACCTGGATGAGCTCCAGCTGGAGATGGAGGACTCAAAAGCCACACCCAGTGTCCAGTG	1504
Qy	211		sSerProThrProGlyProPheLyProCysGluTyrlieuPheGluSerTrpGlyIleAr	231
Db	1505		TAGCCCTACTCCAGGCCCTTCAAGCCCTGTGAGTAGCTCTTTTGAAGCTTGGGCATCCG	1564
Qy	231		gLeuAlaValTrpAlaIleValLeuLeuSerValLeuCysAasnGlyLeuValLeuLeuTh	251
Db	1565		CCTGGCCGTGTGGGCCATCGHTTGTCTCCGTGCTCTGCNAATGGACTGGTGTGCTGCAC	1624
Qy	251		rValPheAlaGlyGlyProAlaProLeuProProValLyAsPheValValGlyAlaIleAl	271
Db	1625		CGTGTTCGTGGCGGGCTGTCCCCCTGCCCGGTCAAGTTTGTGTAGTGTGGATTGC	1684
Qy	271		agLyAlaAasnThrLeuThrGlyIleSerCysGlyLeuLeuAlaSerValAlaLeuTh	291
Db	1685		AGGCGCCCAACACTTGACTGGCAVTTCTGTGGCTTCTAGCCCTCAGTCGATGCCCTGCAC	1744
Qy	291		rPheGlyGlnPheSerGluTyrglyAlaArgTrpGluThrGlyLeuGlyCyAsArgAlaTh	311
Db	1745		CTTTGGTCAGTTCTCTGAGTACGAGCGCCGTGGAGACGGGTAGGCTGCCGGGCCAC	1804
Qy	311		rgLyPheLeuAlaValLeuGlySerGluAlaSerValLeuLeuLeuThrLeuAlaIaVa	331
Db	1805		TGGCTTCTTGGCAGTACTTGGTTCGGAGGCATCGGTGCTGCTGCTACTCTGGCCCGCAGT	1864
Qy	331		lGlnCySerValSerValSerCysValArgAlaTyrglyLySerProSerLeuGlySe	351
Db	1865		GCAGTCAGCGCTCTCCGTCTCTGTCTCGGGCTATGGGAAGTCCCCCTCCCTGGGCAG	1924
Qy	351		rValArgAlaGlyValLeuGlyCysLeuAlaLeuAlaGlyLeuAlaIaIaLeuProLe	371
Db	1925		CGTTCGACAGGGGTCTTAGGCTGCCCTGGCACTGGCAGGGCTGCCCGCGCTGCCCT	1984
Qy	371		uAlaSerValGlyGluTyrglyAlaSerProLeuCysLeuProTyrrAlaProGluGl	391
Db	1985		GGCCTCAGTGGAGAAATACGGGGCTCCCCACTCTGCTGCCCTACGGGCCACTTGAGGG	2044
Qy	391		yGlnProAlaAlaLeuGlyPheThrValAlaLeuValMetMetAasnSerPheCysPheLe	411
Db	2045		TCAGCCAGCAGCCCTGGGCTTCAACGTGGCCCTGGTGATGATGAATCTCTTGTGTTCT	2104
Qy	411		uValValAlaGlyAlaTyrlleLyLeuTyrcysAspLeuProArgLyAspPheGluAl	431
Db	2105		GGTGTGGCCGGTGCCTTACATCAAACTGTACTGTGACCTCCCGCGGGCGACTTTGAGCG	2164
Qy	431		aValTrpAspCysAlaMetValArgHisValAlaTrpLeuIlePheAlaAspGlyLeuLe	451
Db	2165		CGTGTGGGACTGCGCCATGTGTAGGCACGTGGCCTTGCTCATCTTCGACAGCGGGCTCT	2224
Qy	451		uTyrcysProValAlaPheLeuSerPheAlaSerMetLeuGlyLeuPheProValThrPr	471

Db	2225	CTACTGTCGGCTGCTCTCTCAAGTTTGCTTCATGCTGGGCTCTTCCTCTGTACGCC	2284
Qy	471	oGluAlaValIySerValLeuLeuValValLeuProLeuProAlaCysLeuAsnProLe	491
Db	2285	CGAGGCGTCAAGTCTGCTGCTGGTGGTGTGCTGCCCTGCTTCCCTTCAACCCACT	2344
Qy	491	uLeuTyZLeuLeuPheAsnProHisPheArgAspAspLeuArgArgLeuArgProArgAl	511
Db	2345	GCTGTACTGCTCTTCAACCCCACTTCCGGGATGACCTTTCGGCGGCTTCGGCCCCGCGC	2404
Qy	511	aGlyAspSerGlyProLeuAlaTyRAlaAlaAaGlyGluLeuGluIySerSerCysAs	531
Db	2405	AGGGGACTCAGGGGCCCCCTAGCCTATGCTGCGGGCGGGAGCTGGAGAGAGCTCTCTGTGA	2464
Qy	531	pSerThrGlnAlaLeuValAlaPheSerAspValAspLeuIleLeuGluAlaSerGluAl	551
Db	2465	TTCTACCAAGGCCCTGTAGCCTTCTCTGATGTGGATCTCATTTCTGGAGCTTCTGNAGC	2524
Qy	551	aGlyArgProProGlyLeuGluThrTyRglyPheProSerValThrLeuIleSerCysGl	571
Db	2525	TGGCGGCCCCCTGGGCTGGAGACCTATGCTTCCCTCAGTCAGCCCTCATCTCTCTGTCA	2584
Qy	571	nGlnProGlyAlaProArgLeuGluGlySerHisCysValGluProGluClyAsnHisPh	591
Db	2585	GCAGCCAGGGGCCCCCAGGCTGAGGGCAGCCCAATTGTGTAGAGCCAGAGGGGAACCACTT	2644
Qy	591	eGlyAsnProGlnProSerMetAspGlyGluLeuLeuLeuArgAlaGluGlySerThrPr	611
Db	2645	TGGGAACCCCAACCTCTCATGTATGGAGAACTGTGCTGAGGGCAGAGGAGTCTACGCC	2704
Qy	611	oAlaGlyGlyGlyLeuSerGlyGlyGlyGlyPheGlnProSerGlyLeuAlaPheAlaSe	631
Db	2705	AGCAGGTGGAGGCTGTTCAGGGGGTGGCGGCTTTCAGCCCTCTGGCTTGGCCTTTGCTTC	2764
Qy	631	rHisVal 633	
Db	2765	ACAGTG 2771	
RESULT 15			
LOCUS	CQ840811	3306 bp	linear
DEFINITION	Sequence 1 from Patent WO2004059818.		
ACCESSION	CQ840811		
VERSION	CQ840811.1		
KEYWORDS	GI:50838402		
SOURCE	Homo sapiens (human)		
ORGANISM	Homo sapiens		
REFERENCE	1		
AUTHORS	Muda, M.		
TITLE	Spliced variants of Igr6		
JOURNAL	Patent: WO 2004059818-A 1 15-JUL-2004;		
FEATURES	Applied Research Systems ARS Holding N.V. (AN)		
source	Location/Qualifiers		
	1..3306		
	/organism="Homo sapiens"		
	/mol_type="unassigned DNA"		
	/db_xref="taxon:9606"		
	47..2794		
	/notes="unnamed protein product"		
	/codon_start=1		
	/protein_id="CAH05693.1"		
	/db_xref="GI:50838403"		
	/translation="MGRPLRLTVCVQSIISARDLSMMNLTELQPLGFHHLRFLBELR		
	LSGNHLSHIPQWAGFSGLYSLKILMNQNGGIPAEALWELPSQLSLRDANLSLVP		
	ERSFEGSLSLRLHLDNDALTEPVRALNNLPALQAMTILANRISHIPDYAFQNTLS		
	VVLHNNRIQHGLTHSFGSLHNETLDLNYNKLQEPVPIRTGLRQLGPHNNNIK		
	AIPEKFMGNPLLOTTHFYDNPQFVGRSAFOYLPKLHPLSLNGAMDIOEFPDLKCTT		
	LSLEILTRAGIRLLPSGMCQQLPRLRVLESHNQIELPSLHRCOKLSEIGLOHNR		
	WEIGADTFSQLSDLSWNAIRSHPEAFSTLSLVKGLDLDNLQTLTLPGLUGG		
	LMHLKLGMLALSQAFSKDSFPKRLILEVYPAYCCPGYGCASFPGKASQGWAEADLHL		
	DDRESSKRPGLGLARQENHYDQDLDELQLEMEDSKPHPSVQCPTGPGFKPCCEYLFE		
CDS			

SWGRLAVAIWLVLSVLCNGLVLLTVFAGGVPVLPVPKVFVGAAGANTLTGISGCLL
 ASVDALTQGFSEYGEARWETGLGRATGFLAVLGSEASVLLLTAAVQCSVSVSVRA
 YKSPSLGSVAGVLGCLALAGLAAALPLASVGEASPLCLPYAPBEGQPAALGFTV
 ALVMNSFCFLVAGAYTKLYCDLPRGDFEAVWDCAMVRHVAVLIFADGLLYCPVFL
 SPYAMGLGFPVTPPEAVKSVLLVJLPLACILNPLLYLLFNPHFRDLDRLRLRPRAGDSGP
 LAYAAAGLEKSSCDSTQALVAFSDVLLLEASAGRPGLFTYGFPSVTLISQQPG
 APRLEGSCHVPEGNHFNQPPQPSMDGELLRLRAEGSTPAGGGLSGGGFQPSGLAFASH
 V"

ORIGIN

Alignment Scores:

Pred. No.: 2,06e-207 Length: 3306
 Score: 3041.50 Matches: 601
 Percent Similarity: 92.46% Conservative: 12
 Best Local Similarity: 90.65% Mismatches: 20
 Query Match: 91.97% Indels: 32
 DB: 6 Gaps: 4

US-10-664-667-5 (1-633) x Q0840811 (1-3306)

QY 1 AenThrThrHisTyr----- 5
 DB 807 AACTCCACACACTATCTCTGAATGTGCGCATCGACATCCAGAGTTTCCAGATCTCAAG 866
 QY 6 ---ArgGluSerTrpTyrAla-----CysArgTyrArg 15
 DB 867 GCACCACAGCTGGAGATCTCGACCTCGACCCGCGCAGGATCCGGCTGCTCCCATCGG 926
 QY 16 -----SergIlyleProGlySer-ThrHisAlaSerValGluArgSergln---G1 31
 DB 927 GGATGTGCCAACAGCTGCCAGGCTCCGAGTCTCGAACTGTCTCACAATCAAAATTGAGG 986
 QY 31 YleuSerLeuProAlaHisProAlaSerLeuAlaAlaLeuAlaAlaSerAsnThrAl 51
 DB 987 AGCTGCCAGCTGCACAGGTGTGAGAAATTGGAGGAAA-TCGGGCTCCACACACACGCG 1045
 QY 51 aSerGlyLysLeuGlu***AspThrPheSerGlnLeuSerSerLeuGlnAlaLeuAspLe 71
 DB 1046 ATCTGGGAATTGGAGCT-GACACCTTCAGCAGCTGAGCTCCCTGCAAGCCCTGGATCT 1104
 QY 71 uSerTrpAsnAlaIleArgSerIleHisProGluAlaPheSerThrLeuHisSerLeuVa 91
 DB 1105 TAGCTGGAAACCATCCGGTCCATCCACCCGAGGCTTCTCCACCCCTGCACCTCCCTGGT 1164
 QY 91 llyLeuAspLeuThrAspAsnGlnLeuThrThrLeuProLeuAlaGlyLeuGlyLe 111
 DB 1165 CAAGCTGCACCTGACAGAACCAAGCTGACACACTGCCCTTGGCTGGACTTGGGGGCTT 1224
 QY 111 uMetHisLeuLysLeuLysGlyAsnLeuAlaLeuSerGlnAlaPheSerLysAspSerPh 131
 DB 1225 GATGCATCTGAAGCTCAAGGGGAACCTTGCTCTCCAGGCTTCTCCAGGACAGATT 1284
 QY 131 eProLysLeuArgIleLeuGluValProTyrAlaTyrGlnCysCysProTyrGlyMetCy 151
 DB 1285 CCAAAACTGAGGATCCTGGAGTGCCCTTATGCTTACCAGTGCTGCTCCCTATGGATGTG 1344
 QY 151 sAlaSerPhePheLysAlaSerGlyGlnTrpGluAlaGluAspLeuHisLeuAspAspG1 171
 DB 1345 TGCCAGCTTCTCAAGGCTCTGGGAGCTGGAGGCTGAAGACCTTCACTTGTATGATGA 1404
 QY 171 uGluSerSerLysArgProLeuGlyLeuAlaArgGlnAlaGluAsnHisTyrAspG1 191
 DB 1405 GGAGTCTTCAAAAGGCCCTTGGGCTCTTGGCCAGACAGAGAGAACCTATGACCA 1464
 QY 191 nAspLeuAspGluLeuGlnLeuGluMetGluAspSerLysProHisProSerValGlnCy 211
 DB 1465 GGACCTGATGAGCTCCAGCTGGAGATGGAGACTCAAAAGCCACACCCAGCTGCCAGTG 1524
 QY 211 sSerProThrProGlyProPheLysProCysGluTyrLeuPheGluSerTrpGlyIleAr 231
 DB 1525 TAGCCCTACTCCAGGCCCTTCAAGCCCTGTGAGTACCTCTTTGAAAGCTGGGGCATCCG 1584
 QY 231 gLeuAlaValTrpAlaIleValLeuLeuSerValLeuLysAsnGlyLeuValLeuLeuTh 251

DB 1585 CCTGGCGTGTGGGCCATCGTGTGCTCTCGTCTCTGCAATGACTGTGTGCTGAC 1644
 QY 251 rValPheAlaGlyClyProAlaProLeuProProValIysPheValValGlyAlaIleAl 271
 DB 1645 CGTGTTCGCTGGCGGGCTGTCCCCCTGCCCGCGTCAAGTTTGTGGTAGGTGGATGTC 1704
 QY 271 aGlyAlaAsnThrLeuThrGlyIleSerCysGlyLeuLeuAlaSerValAspAlaLeuTh 291
 DB 1705 AGGCGCCAAACCTTGACTGGCATTTCTGTGGCTTCTAGCCTCAGTCGATGCCGTGAC 1764
 QY 291 rPheGlyGlnPheSerGluTyrGlyAlaArgTrpGluThrGlyLeuGlyCysArgAlaTh 311
 DB 1765 CTTTGGTCAGTTCTCTCAGTAGCGAGCCGCTGGAGAGCGGGCTAGGCTCCCGGGCCAC 1824
 QY 311 rGlyPheLeuAlaValLeuGlySerGluAlaSerValLeuLeuLeuThrLeuAlaAlaVa 331
 DB 1825 TGGCTTCTCGCAGTACTTGGTCTGGAGGATCGGTGCTGCTGCTCACTCTGGCGCGCAGT 1884
 QY 331 lGlnCysSerValSerValSerCysValArgAlaTyrGlyLysSerProSerLeuGlySe 351
 DB 1885 GCAGTGAGGCTTCCGCTCTCTGTGTCGGGGCTATGGAGAGTCCCTCCCTCGGGCAG 1944
 QY 351 rValArgAlaGlyValLeuGlyCysLeuAlaLeuAlaGlyLeuAlaAlaLeuProLe 371
 DB 1945 CGTTTCGAGCAGGGGTCTTAGGCTGCCTGGCAGTGGCAGGGCTGGCGCGGCTGCCCT 2004
 QY 371 uAlaSerValGlyClyTyrGlyAlaSerProLeuCysLeuProTyrAlaProGluG1 391
 DB 2005 GGCTTCAGTGGAGATACGGGGCTCCCCACTCTGCTGCCCTACGCCGCCACCTGAGGG 2064
 QY 391 YGlnProAlaAlaLeuGlyPheThrValAlaLeuValMetMetAsnSerPheCysPheLe 411
 DB 2065 TCAGCCAGCAGCCCTGGGCTTACCGTGGCCCTGGTGATGATGAACTCTCTTGTTCCT 2124
 QY 411 uValValAlaGlyAlaTyrIleLysLeuTyrCysAspLeuProArgGlyAspPheGluAl 431
 DB 2125 GGTCTGCGCGGTGCTTACATCAACTGTACTGTGACTGCCGCGGGCGAGCTTTGAGGC 2184
 QY 431 aValTrpAspCysAlaMetValArgHisValAlaTrpIlePheAlaAspGlyLeuLe 451
 DB 2185 CGTGTGGAGCTGGCGCCATGGTAGGACAGTGGCTGCTCATCTTCGACAGCGGGCTCT 2244
 QY 451 uTyrCysProValAlaPheLeuSerPheAlaSerMetLeuGlyLeuPheProValThrPr 471
 DB 2245 CTACTGTCCGTGGCTTCTCAGCTTGTGCTCATGCTGGGCTCTTCCCTGTACGCC 2304
 QY 471 oGluAlaValLysSerValLeuLeuValLeuProLeuProAlaCysLeuAsnProLe 491
 DB 2305 CGAGGCGTCAAGTCTGCTGCTGTGTGGTGTGCTGCCCTGCTGCTCAACCCACT 2364
 QY 491 uLeuTyrLeuLeuPheAsnProHisPheArgAspAspLeuArgArgLeuArgProArgAl 511
 DB 2365 GCTGTACTGCTCTTCAACCCCACTTCGCGGATGAGCTTTCGGCGGCTTCGGCCCCGCGC 2424
 QY 511 aGlyAspSerGlyProLeuAlaTyrAlaAlaGlyLysLeuGlyLysSerSerCysAs 531
 DB 2425 AGGGAGCTCAGGGCCCCCTAGCTTATGCTGCGCCCGGGAGCTGGAGAAGAGCTCTGTGA 2484
 QY 531 pSerThrGlnAlaLeuValAlaPheSerAspValAspLeuLeuGluAlaSerGluAl 551
 DB 2485 TTCTACCCAGGCCCTGTGAGCTTCTCTGATGTGATCTCATCTTGGAAAGCTTCTGAAGC 2544
 QY 551 aGlyArgProProGlyLeuGluThrTyrGlyPheProSerValThrIleLysSerCysG1 571
 DB 2545 TGGCGGCCCTCCCTGGAGACCTATGGCTTCCCTCAGTGAGTCCCTCATCTCTCTGTCA 2604
 QY 571 nGlnProGlyAlaProArgLeuGluGlySerHisCysValGluProGluGlyAsnHisPh 591
 DB 2605 GCAGCCAGGGGCCCTCAGGCTGGAGGCGAGCCATTGTGTAGAGCCAGAGGGGAACCATCT 2664
 QY 591 eGlyAsnProGlnProSerMetAspGlyGluLeuLeuLeuArgAlaGluGlySerThrPr 611

Db	2665	TGGGAACCCCAACCTCCATGGATGGAGAACTGCTGTGAGGGCAGAGGGATCTACGCC	2724
Qy	611	oAlaGlyGlyGlyLeuSerGlyGlyGlyPheGlnProSerGlyLeuAlaPheAlase	631
Db	2725	AGCAGGTGGAGGCTTGTCAGGGGGTGGGGCTTTCAGCCCTCTGGCTTGGCCTTTGCTTC	2784
Qy	631	rHisVal	633
Db	2785	ACACGTG	2791

Search completed: April 14, 2005, 20:48:36
Job time : 5075 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: April 14, 2005, 12:30:08 ; Search time 613 Seconds
(without alignments)
6112.881 Million cell updates/sec

Title: US-10-664-667-5

Perfect score: 3307

Sequence: 1 NTHYRESWACRYVRSGIPG.....GGLSGGGFGPQGLAFASHV 633

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 4390206 seqs, 2959870667 residues

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlh
-Q=/cgn2_1/USPTO.spool/US10664667/runat_13042005_184539_7252/app_query.fasta_1.775
-DB=Geneseq_16Dec04 -QMT=faastep -SUFFIX=p2n.rng -MINMATCH=0.1 -LOOPCL=0
-LOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=pco -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000
-USER=US10664667_QCGN_1_1_352 @runat_13042005_184539_7252 -NCPU=6 -ICPU=3
-NO MMAP -LARGQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

N.Geneseq_16Dec04:*
1: Geneseqn1980s:*
2: Geneseqn1990s:*
3: Geneseqn2000s:*
4: Geneseqn2001as:*
5: Geneseqn2001bs:*
6: Geneseqn2002as:*
7: Geneseqn2002bs:*
8: Geneseqn2003as:*
9: Geneseqn2003bs:*
10: Geneseqn2003cs:*
11: Geneseqn2003ds:*
12: Geneseqn2004as:*
13: Geneseqn2004bs:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3305	99.9	1899	AAI67923	AAI67923 Human LGR
2	3305	99.9	1899	11 ADN02244	Adn02244 Human par
3	3305	99.9	1899	12 ADK19409	Adk19409 ORF of CD
4	3305	99.9	2486	AAI67922	AAI67922 Human LGR
5	3305	99.9	2486	11 ADN02242	Adn02242 Human par

6	3305	99.9	2486	12 ADK19407	Adk19407 cDNA enco
7	3150	95.3	2322	5 AAF98722	AAf98722 Human lat
8	3140	95.0	3451	12 ADQ76775	Adq76775 Human LGR
9	3045.5	92.1	2208	6 AAI67925	AAI67925 Human LGR
10	3045.5	92.1	2208	11 ADN02247	Adn02247 Human par
11	3045.5	92.1	2208	12 ADK19412	Adk19412 ORF of pa
12	3045.5	92.1	2711	6 AAI67924	AAI67924 Partial c
13	3045.5	92.1	2711	11 ADN02245	Adn02245 Human par
14	3045.5	92.1	2711	12 ADK19410	Adk19410 Partial c
15	3045.5	92.1	2786	10 ADK16710	Adc16710 Human G-p
16	3045.5	92.1	2901	6 AAI67927	AAI67927 Human LGR
17	3045.5	92.1	2901	10 ADK16695	Adc16695 cDNA with
18	3045.5	92.1	2901	10 ADK16692	Adc16692 cDNA enco
19	3045.5	92.1	2901	11 ADN02250	Adn02250 Human ful
20	3045.5	92.1	2901	12 ADK19415	Adk19415 ORF of fu
21	3045.5	92.1	3000	6 AAS98067	AAS98067 Human DNA
22	3045.5	92.1	3042	10 ADK16701	Adc16701 cDNA with
23	3045.5	92.1	3042	10 ADK16698	Adc16698 cDNA enco
24	3045.5	92.1	3119	8 ABZ42549	ABZ42549 Human G-p
25	3045.5	92.1	3325	10 ADD89048	Add89048 Encoding
26	3045.5	92.1	3429	6 AAD37674	AAd37674 Human G-p
27	3045.5	92.1	3438	10 ABT31938	ABt31938 Human bre
28	3045.5	92.1	3443	10 ADD89090	Add89090 Encoding
29	3045.5	92.1	3492	6 AAI67926	AAI67926 Human LGR
30	3045.5	92.1	3492	11 ADN02248	Adn02248 Human ful
31	3045.5	92.1	3492	12 ADK19413	Adk19413 Full leng
32	3045.5	92.1	3492	12 ADQ15073	Adq15073 Human can
33	3041.5	92.0	2901	10 ADK16694	Adc16694 cDNA enco
34	3041.5	92.0	2901	10 ADK16696	Adc16696 cDNA enco
35	3041.5	92.0	2964	10 ADI21981	Adi21981 Novel hum
36	3041.5	92.0	3042	10 ADK16702	Adc16702 cDNA with
37	3041.5	92.0	3042	10 ADK16700	Adc16700 cDNA enco
38	3041.5	92.0	3273	11 ADN39801	Adn39801 Cancer/an
39	3041.5	92.0	3273	13 ADK46630	Adk46630 Cancer-as
40	3041.5	92.0	3306	12 ADQ76773	Adq76773 Human LGR
41	3041.5	92.0	3618	10 ADF70543	Adf70543 Orphan re
42	3040.5	91.9	2487	6 ADH50804	Adh50804 Human G-p
43	3040	91.9	2988	6 ABZ11841	ABz11841 Human pol
44	3040	91.9	2988	12 ADM44359	Adm44359 Novel hum
45	3038.5	91.9	3349	6 ADH50807	Adh50807 Human G-p

ALIGNMENTS

RESULT 1	AAI67923	AAI67923 standard; cDNA; 1899 BP.
ID	AAI67923	standard; cDNA; 1899 BP.
XX	AAI67923;	
AC	AAI67923;	
XX	13-MAR-2002	(first entry)
DT	Human LGR6	polypeptide coding sequence (clone fahr).
DE	Human LGR6	polypeptide coding sequence (clone fahr).
XX	Large G-protein coupled receptor 6; LGR6; GPCR; immunomodulator;	
KW	anti-HIV; cytostatic; neurotropic; neuroprotective; antiparkinsonian;	
KW	anticonvulsant; neuroleptic; antidepressant; tranquilizer; antimigraine;	
KW	osteopathic; antiarteriosclerotic; vasotropic; hypotensive; cardiant;	
KW	antiarrhythmic; anorectic; gene therapy; human; 55.	
XX	Homo sapiens.	
OS	Key	Location/Qualifiers
FH	CDS	1..1899
FT		/*tag= a
FT		/product= "LGR6 polypeptide"
FT		/transl_except= (pos: 169..171, aa: Xaa)
FT		/note= "Xaa = unknown"
XX	WO200185768-A2.	
PN	15-NOV-2001.	
XX		
PD		
XX		

```
PF 08-MAY-2001; 2001WO-US015002.
XX
FR 08-MAY-2000; 2000US-00566588.
XX
PA (MILL-) MILLENNIUM PHARM INC.
XX
PI Gu W;
XX
XX WPI; 2002-055584/07.
DR P-PSDB; AAG66139.
XX
XX Novel isolated large G-protein coupled receptor 6 polypeptide, useful for
PT diagnosing and treating weight disorder, metabolic disorder, central
PT nervous system disorder, endocrine disorder and cardiovascular disorder.
XX
XX Example 1; Fig 4; 198pp; English.
XX
XX The invention provides isolated large G-protein coupled receptor 6 (LGR6)
CC polypeptides and polynucleotides encoding them. The LGR6 polypeptides and
CC polynucleotides are useful as targets for developing modulating agents
CC that regulate a variety of cellular process, e.g. neural and endocrine
CC processes, as well as thermogenesis. They are useful for developing novel
CC diagnostic and therapeutic agents for LGR6 associated disorders such as
CC weight disorders (anorexia, obesity), eating disorders, acquired
CC immunodeficiency syndrome (AIDS)-related wasting, cancer-related wasting,
CC metabolic disorders, central nervous system disorders (e.g. Alzheimer's
CC disease, Parkinson's disease, Huntington's disease, Gilles de la
CC Tourette's syndrome, multiple sclerosis, amyotrophic lateral sclerosis,
CC epilepsy, Jakob-Creutzfeldt disease, depression, schizophrenia,
CC obsessive-compulsive disorder, anxiety, panic disorder, migraine),
CC endocrine disorders (hypo and hyperthyroidism, dwarfism, gigantism,
CC acromegaly), and cardiovascular disorders (atherosclerosis, stenosis,
CC hypertension, myocardial infarction, arrhythmia). The LGR6 polypeptides,
CC polynucleotides and antibodies are useful in screening assays, detection
CC assays (e.g., chromosomal mapping, tissue typing, forensic biology),
CC predictive medicine (e.g. diagnostic assays, prognostic assays,
CC monitoring clinical trials and pharmacogenomics), and in methods of
CC treatment (e.g. therapeutic and prophylactic). The present sequence
XX represents a human LGR6 polypeptide coding sequence
XX
SQ Sequence 1899 BP; 302 A; 624 C; 563 G; 409 T; 0 U; 1 Other;

Alignment Scores:
Pred. No.: 8,45e-259 Length: 1899
Score: 3305.00 Matches: 632
Percent Similarity: 99.84% Conservative: 0
Best Local Similarity: 99.84% Mismatches: 1
Query Match: 99.94% Indels: 0
DB: 6 Gaps: 0

US-10-664-667-5 (1-633) x AAI67923 (1-1899)
Qy 1 AsnThrThrHisTyrArgGluSerTrpTyrAlaCysArgTyrArgSerGlyLeuProGly 20
Dy 1 AATACGACTCACTATAGGAAAGCTGTGTACGCTCGCAGGTACCGTCCGGAAATCCCGGG 60
Qy 21 SerThrHisAlaSerValGluArgSerGlnGlyLeuSerLeuProAlaHisProAlaSer 40
Dy 61 TCGACCCACGCGTCCGTGTGAGCGGAGCCAGGCTGTAGCGCTCGCGGCTCATCCAGCCTCT 120
Qy 41 LeuAlaAlaLeuAlaAlaSerAsnThrThrAlaSerGlyLeuGlu***AspThrPhe 60
Dy 121 CTTCGTCGCTAGCGGCTCCCAACACACACCGCATCTGGGAAATGGAGCTNGACACCTTC 180
Qy 61 SerGlnLeuSerSerLeuGlnAlaLeuAspLeuSerTrpAsnAlaIleArgSerIleHis 80
Dy 181 AGCCAGCTGAGTCCCTCGAAGCCCTGGATCTTAGTGTGAACCGCATCCGCTCATCCAC 240
Qy 81 ProGluAlaPheSerThrLeuHisSerLeuValLysLeuAspLeuThrAspAsnGlnLeu 100
Dy 241 CCTGAGGCTTCTCCACCCCTCACTCCCTGCTGAAGCTGGACCTGCACAGACACAGCTG 300
Qy 101 ThrThrLeuProLeuAlaGlyLeuGlyLeuMetHisLeuLysLeuLysGlyAsnLeu 120
301 ACCACACTGCCCTGGCTGGACTTGGGGCTTGATGATCTGAAGCTCAAGGGACCTT 360
121 AlaLeuSerGlnAlaPheSerLysAspSerPheProLysLeuArgIleLeuGluValPro 140
361 GCTCTCTCCAGGCTTCTCCAAAGACAGATTTCCAAAACCTGAGGATCTCGAGGTGCT 420
141 TyrAlaTyrGlnCysCysProTyrGlyMetCysAlaSerPhePheLeuAlaSerGlyGln 160
421 TATGCCTACCAGTCTCTCCCTATGGGATGTGTCCAGCTTCTTCAAGGCTCTCGGGCAG 480
161 TrpGluAlaGluAspLeuHisLeuAspAspGluSerSerLysArgProLeuGlyLeu 180
481 TGGGAGCTGAGACCTTCACTTGATGATGAGGAGTCTTCAAAAGGCCCTCGGGCTC 540
181 LeuAlaArgGlnAlaGluAsnHisTyrAspGlnAspLeuAspGluLeuGluMet 200
541 CTTGCCACAGACAGAGAACCACTATGACAGGACCTGGATGAGCTCCAGCTGGAGATG 600
201 GluAspSerLysProHisProSerValGlnCysSerProThrProGlyProPheLeuValPro 220
601 GAGGACTCAAAGCCACACCCAGTGTCCAGTGTAGCCCTACTCCAGGCCCTTCAAGCCC 660
221 CysGluTyrLeuPheGluSerTrpGlyIleArgLeuAlaValTrpAlaIleValLeuLeu 240
661 TGTGAGTACCTCTTTGAAAGCTGGGGCATCCGCTGCGGTGGGGCATCGTGTGCTC 720
241 SerValLeuCysAsnGlyLeuValLeuThrValPheAlaGlyGlyProAlaProLeu 260
721 TCCGTGCTCTGCAATGAGCTGGTGTCTGCTGACCGTGTGGGGCTTCCGCCCTG 780
261 ProProValLysPheValValGlyAlaIleAlaGlyAlaAsnThrLeuThrGlyLeuSer 280
781 CCCCCGCTCAAGTTTGGTAGTGCGATTCGAGGCCCAACACCTTGTGACTGGCATTTCC 840
281 CysGlyLeuLeuAlaSerValAspAlaLeuThrPheGlyGlnPheSerGluTyrGlyAla 300
841 TGTGGCTTCTAGCTCAGTCGATGCCCTGACCTTTGGTCAGTCTCTGTAGTACGGAGCC 900
301 ArgTrpGluThrGlyLeuGlyCysArgAlaThrGlyPheLeuAlaValLeuGlySerGlu 320
901 CGCTGGGAGACGGGGCTAGGCTGCCGGGCCACTGGCTTCCTGGCAGTACTTGGGTCGAG 960
321 AlaSerValLeuLeuLeuThrLeuAlaAlaValGlnCysSerValSerValSerCysVal 340
961 GCATCGGTGTGTGCTCACTCTGGCCGCGAGTGCAGTGCAGCGTCTCCGCTCTCTGTGTC 1020
341 ArgAlaTyrGlyLysSerProSerLeuGlySerValArgAlaGlyValLeuGlyCysLeu 360
1021 CGGGCTATGGGAAGTCCCTCCCTGGGCGAGGCTTCGAGCAGGGGTCTTAGGCTGCTG 1080
361 AlaLeuAlaGlyLeuAlaAlaLeuProLeuAlaSerValGlyGlyTyrGlyAlaSer 380
1081 GCATGGCAGGGCTGGCCGCGCACCTGCCCTCGGCTCAGTGGGAGAATACGGGGCTCC 1140
381 ProLeuCysLeuProTyrAlaProProGluGlyGlnProAlaLeuGlyPheThrVal 400
1141 CCATCTCGCTGCTGCTACCGCCGCGCACCTGAGGGTCAGCCAGCGCCCTTCCCGGTG 1200
401 AlaLeuValMetMetAsnSerPheCysPheLeuValValAlaGlyAlaTyrIleLysLeu 420
1201 GCCTGTGTGATGATGAATCTCTTCTGTGCTGGTGGCGGCTGCTATACAAACTG 1260
421 TyrCysAspLeuProArgGlyAspPheGluAlaValTrpAspCysAlaMetValArgHis 440
1261 TACTGTGACCTGCCGCGGGCGACTTTAGGCGCGTGTGGGACTGCGCGCATGGTGAGGCAC 1320
441 ValAlaTrpLeuIlePheAlaAspGlyLeuLeuTyrCysProValAlaPheLeuSerPhe 460
1321 GTGGCTTGGCTCACTTCCTGCGAGACGGGCTCTCTTACTGTCCCGGGGCTTCTCAGCTTC 1380
461 AlaSerMetLeuGlyLeuPheProValThrProGluAlaValLysSerValLeuVal 480
```



```
Qy 141 TyrAlaTyrGlnCysCysProTyrGlyMetCysAlaSerPhePheLysAlaSerGlyGln 160
Db 421 TATGCTTACCAGTGTGTCCTTATGGATGTGTCCAGCTTCTTCAAGGCTCTGGGGCAG 480
Qy 161 TrpGluAlaGluAspLeuHisLeuAspAspGluGluSerSerLysAspTrpLeuGlyLeu 180
Db 481 TGGGAGGCTGAAGACCTTCACTTGATGATGAGGAGTCTTCAAAAAGGCCCTTGGGCTC 540
Qy 181 LeuAlaArgGlnAlaGluAsnHisTyrAspGlnAspLeuAspGluLeuGlnLeuGluMet 200
Db 541 CTTGCCAGACAAGCAGAGAACCACTATGACCAGGACCTGGATGAGCTCCAGCTGGAGATG 600
Qy 201 GluAspSerLysProHisProSerValGlnCysSerProThrProGlyProPheLysPro 220
Db 601 GAGGACTCAAAAGCCACACCCAGTGTCCAGTGTAGCCCTACTCCAGGCCCCCTTCAAGCCC 660
Qy 221 CysGluTyrLeuPheGluSerTrpGlyIleArgLeuAlaValTrpAlaIleValLeuLeu 240
Db 661 TGTGAGTACCTCTTTGAAAGCTGGGGCATCCGCTGGCCGTTGGGCCATCGTGTGCTC 720
Qy 241 SerValLeuCysAsnGlyLeuValLeuLeuThrValPheAlaGlyGlyProAlaProLeu 260
Db 721 TCCGTGCTCTGCAATGGACTGGTGTGCTGACCGTGTTCGCTGGCGGCCCTGCCCCCTG 780
Qy 261 ProProValLysPheValValGlyAlaIleAlaGlyAlaAsnThrLeuThrGlyLysSer 280
Db 781 CCCCCTGTCAGTTGTGGTGTAGGTGCGATTGCAGCGCCCAACACCTTGACTGGCATTTCC 840
Qy 281 CysGlyLeuLeuAlaSerValAspAlaLeuThrPheGlyGlnPheSerGluTyrGlyAla 300
Db 841 TGTGGCTTCTAGCCTCAGTCGATGCCCTGACCTTTGGTCAGTTCCTGAGTACGGAGCC 900
Qy 301 ArgTrpGluThrGlyLeuGlyCysArgAlaThrGlyPheLeuAlaValLeuGlySerGlu 320
Db 901 CGCTGGGAGACGGGGCTAGGCTGCGGGCCACTGGCTTCCTGGCAGTACTTGGGTCCGAG 960
Qy 321 AlaSerValLeuLeuThrLeuAlaValGlnCysSerValSerValSerCysVal 340
Db 961 GCATCGGGTGTGCTGCTCAGTCTGGCCGACGTGAGTGCAGCGTCCGTTCTCTGTGTC 1020
Qy 341 ArgAlaTyrGlyLysSerProSerLeuGlySerValArgAlaGlyValLeuGlyCysLeu 360
Db 1021 CGGGCCTATGGAGTATCCCTCCCTGGGACGCTTCGAGCAGGGTCTTAGGTGGCTG 1080
Qy 361 AlaLeuAlaGlyLeuAlaAlaLeuProLeuAlaSerValGlyGluTyrGlyAlaSer 380
Db 1081 GCACCTGGCAGGGCTGGCGCGCACCTGCCCTGCGCTCAGTGGGAGAAATACGGGGCCTCC 1140
Qy 381 ProLeuCysLeuProTyrAlaProProGluGlyGlnProAlaAlaLeuGlyPheThrVal 400
Db 1141 CCATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1200
Qy 401 AlaLeuValMetMetAsnSerPheCysPheLeuValValAlaGlyAlaTyrIleLysLeu 420
Db 1201 GCCCTGGTGATGATGAATCTCTCTGTTCTGTTGCTGCTGCTGCTGCTGCTGCTGCTG 1260
Qy 421 TyrCysAspLeuProArgGlyAspPheGluAlaValTrpAspCysAlaMetValArgHis 440
Db 1261 TACTGTGACCTGCGCGGGGCACTTTGAGCGCGTGTGGGACTGCGCATGGTGAGGCAC 1320
Qy 441 ValAlaTrpLeuIlePheAlaAspGlyLeuLeuTyrCysProValAlaPheLeuSerPhe 460
Db 1321 GTGGCCCTGGCTCATCTTCGACAGCGGGCTCTCTACTGTCCCGTGGGCTTCTCTCAGCTTC 1380
Qy 461 AlaSerMetLeuGlyLeuPheProValThrProGluAlaValLysSerValLeuLeuVal 480
Db 1381 GCCTCCATGTGGGCTCTTCCCTGTGTCAGGCCCGGCGCTCAAGTCTGTCTCTGCTGGTG 1440
Qy 481 ValLeuProLeuProAlaCysLeuAsnProLeuLeuTyrLeuLeuPheAsnProHisPhe 500
Db 1441 GTGCTGCCCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1500
```

```
Qy 501 ArgAspAspLeuArgArgLeuArgProArgAlaGlyAspSerGlyProLeuAlaTyrAla 520
Db 1501 CGGATGACCTTCGGCGGCTTCGGCCCCGCGCAGGGGACTCAGGGCCCCCTAGCTATGCT 1560
Qy 521 AlaAlaGlyGluLeuGluLysSerSerCysAspSerThrGlnAlaLeuValAlaPheSer 540
Db 1561 GCGCCCGGGAGCTGGAGAAGAGCTCCTGTGATTCTACCCAGGCCCTGGTAGCCTTCTCT 1620
Qy 541 AspValAspLeuLeuLeuGluAlaSerGluAlaGlyArgProProGlyLeuGluThrTyr 560
Db 1621 GATGTGATCTCATTTCTGGAGCTTCTTGAAGCTGGCGGCCCTTGGCTGGAGACCTAT 1680
Qy 561 GlyPheProSerValThrLeuLeuSerCysGlnGlnProGlyAlaProArgLeuGluGly 580
Db 1681 GGCTTCCCTCAGTACCTCTCATCTCTGTGTCAGCAGCAGCGGGGCCCTGAGGGGC 1740
Qy 581 SerHisCysValGluProGluGlyAsnHisPheGlyAsnProGlnProSerMetAspGly 600
Db 1741 AGCCATTGTGTAGAGCCAGAGGGGACCACTTTGGGAACCCCAACCTCCATGATGGA 1800
Qy 601 GluLeuLeuLeuArgAlaGluGlySerThrProAlaGlyGlyLeuSerGlyGlyGly 620
Db 1801 GAACCTGCTGTGAGGGCAGAGGGATCTACGCCACAGCTGGAGCTTGTACAGGGGTGGC 1860
Qy 621 GlyPheGlnProSerGlyLeuAlaPheAlaSerHisVal 633
Db 1861 GGCTTTCAGCCCTCTGGCTTGGCTTGGCTTTCACACGTG 1899

RESULT 3
ADK19409
ID ADK19409 standard; cDNA; 1899 BP.
XX
AC ADK19409;
XX
DT 17-JUN-2004 (first entry)
XX
DE ORF of cDNA encoding human LGR6 polypeptide #1.
XX
KW G-protein coupled receptor; GPCR; large G-protein coupled receptor 6;
KW LGR6; weight disorder; neural disorder; endocrine disorder;
KW cardiovascular disorder; anorectic; anabolic; immunomodulator;
KW neuroprotective; nootropic; antiparkinsonian; anti-HIV; anticonvulsant;
KW antithyroid; osteopathic; cardiac; vasotropic; hypotensive;
KW antiarteriosclerotic; antiarrhythmic; human; gene; ss.
XX
OS Homo sapiens.
XX
FH Key CDS
FT 1..1899
FT /*tag= a
FT /partial
FT /product= "LGR6 polypeptide #1"
FT /note= "This sequence lacks both start and stop codons"
FT /transal except= (pos:169..171,aa:Xaa)
FT /note= "Xaa= Unknown"
XX
PN US2004058377-A1.
XX
PD 25-MAR-2004.
XX
PF 18-SEP-2003; 2003US-00664667.
XX
PR 06-MAY-1999; 99US-0132896P.
PR 08-MAY-2000; 2000US-00566588.
PR 08-MAY-2001; 2001US-00851595.
XX
PA (MILL-) MILLENNIUM PHARM INC.
XX
PI Gu W;
XX
DR WPI; 2004-268789/25.
DR P-PSDB; ADK19408.
XX
```

PT New large G-protein coupled receptor 6 nucleic acid molecules and
PT polypeptides, useful for diagnosing, preventing or treating diseases
PT associated with aberrant nucleic acid or protein activity, e.g. obesity,
XX anorexia or cachexia.

PS Claim 2; SEQ ID NO 6; 145pp; English.

XX The present invention relates to the isolation of novel members of the G-
CC protein coupled receptor (GPCR) family designated as large G-protein
CC coupled receptor 6 (LGR6), and the polynucleotide sequences encoding
CC them. The invention also discloses LGR6 fusion proteins, antigenic
CC peptides, anti-LGR6 antibodies, recombinant expression vectors, host
CC cells, and non-human transgenic animals in which an LGR6 gene has been
CC introduced or disrupted. The LGR6 polynucleotide sequences, polypeptide
CC sequences, and antibodies are useful for diagnosing, treating or
CC preventing diseases associated with aberrant LGR6 expression or activity,
CC such as weight disorders (e.g. obesity, anorexia nervosa or cachexia),
CC neural disorders (e.g. central nervous system disorder, including
CC Alzheimer's disease, Parkinson's disease, AIDS-related dementia, multiple
CC sclerosis and Huntington's disease), endocrine disorders (e.g.
CC hypothyroidism, hyperthyroidism or acromegaly), or cardiovascular
CC disorders (e.g. ischaemia-reperfusion injury, coronary artery disease,
CC myocardial infarction, arrhythmia, atherosclerosis, hypertension or
CC congestive heart failure). The LGR6 polynucleotide and polypeptide
CC sequences are also useful as targets for developing modulating agents
CC that regulate a variety of cellular processes, e.g. neural and endocrine
CC processes and thermogenesis. They can be used in screening assays (e.g.
CC chromosome mapping, tissue typing or in forensic biology), or in
CC predictive medicine (e.g. diagnostic assays, prognostic assays,
CC monitoring clinical trials or pharmacogenomics). The present sequence
XX encodes a human LGR6 polypeptide.

SQ Sequence 1899 BP; 302 A; 624 C; 563 G; 409 T; 0 U; 1 Other;

Alignment Scores:

Pred. No.: 8,45e-259 Length: 1899
Score: 3305.00 Matches: 632
Percent Similarity: 99.84% Conservative: 0
Best Local Similarity: 99.84% Mismatches: 1
Query Match: 99.94% Indels: 0
DB: 12 Gaps: 0

US-10-664-667-5 (1-633) x ADK19409 (1-1899)

QY 1 AenThrHisTyrArgGluSerTrpTyrAlaCysArgTyrArgSerGlyIleProGly 20
DB 1 AATAGACTACTATAGGGAAGCTGTACCGCTCAGATCCGTCGGAAATCCCGG 60
QY 21 SerThrHisAlaSerValGluArgSerGlnGlyLeuSerLeuProAlaHisProAlaSer 40
DB 61 TCGACCCACGGCTCCGTGGAGCGGAGCCAGGGTCTGAGCTCGCGCTCATCCAGCTCT 120
QY 41 LeuAlaAlaLeuAlaAlaSerHisThrAlaSerGlyLeuGlu***AspThrPhe 60
DB 121 CTGTGTGCCCTAGCGGCCCTCCAAACACCGCATCTGGGAATTTGAGCTNGACACCTTC 180
QY 61 SerGlnLeuSerSerLeuGlnAlaLeuAspLeuSerTrpAsnAlaIleArgSerIleHis 80
DB 181 AGCCAGCTGAGCTCCCTGCAAGCCCTGGATCTTAGCTGGAACGCCATCCGTCCTCCAC 240
QY 81 ProGluAlaPheSerThrLeuHisSerLeuValLysLeuAspLeuThrAspAsnGlnLeu 100
DB 241 CCTGAGGCTTCTCCACCTGACCTCCCTGGTCAAGCTGGACCTGACAGACACACGCTG 300
QY 101 ThrThrLeuProLeuAlaGlyLeuGlyLeuMetHisLeuLysLeuGlyAsnLeu 120
DB 301 ACCACATGCCCTGGCTGGACTTGGGGCTTGATGCATCTGAAGCTCAAGAGGAACCTT 360
QY 121 AlaLeuSerGlnAlaPheSerLysAspSerPheProLysLeuArgIleLeuValPro 140
DB 361 GCTCTCTCCAGGCCCTTCTCCAGGACAGATTTCCCAAACTGAGATCTTGAGGTGCT 420
QY 141 TyrAlaTyrGlnCysCysProTyrGlyMetCysAlaSerPhePheLysAlaSerGlyGln 160

DB 421 TATGCTTACCAAGTGTGTCCCTATGGGATGTGTCCAGCTTCTTCAAGGCCCTCTGGGAG 480
QY 161 TrpGluAlaGluAspLeuHisLeuAspAspGluGluSerSerLysArgProLeuGlyLeu 180
DB 481 TGGGAGGCTGAAGACCTTACCTTGATGATGAGGAGTCTTCAAAAAGGCCCTTGGGCCTC 540
QY 181 LeuAlaArgGlnAlaGluAsnHisTyrAspGlnAspLeuAspGluLeuGlnLeuGluMet 200
DB 541 CTTGCCAGACAGACAGAGAACCACTATGACAGGACCTGGATGAGCTCCAGCTGGAGATG 600
QY 201 GluAspSerLysProHisProSerValGlnCysSerProThrProGlyProPheLysPro 220
DB 601 GAGGACTCAAAAGCCACACCCAGTGTCCAGTGTAGCCCTACTCCAGGCGCCCTTCAAGGCC 660
QY 221 CysGluTyrLeuPheGluSerTrpGlyIleArgLeuAlaValTrpAlaIleValLeuLeu 240
DB 661 TGTGAGTACCTTCTTTGAAGCTGGGGCATCCGCTGGCGGTGGGGCATCTGTGTGCTC 720
QY 241 SerValLeuCysAsnGlyLeuValLeuLeuThrValPheAlaGlyGlyProAlaProLeu 260
DB 721 TCCGTGCTCTGCAATGGACTGTGTCTGTGACCGGTGTCGTGGCGGCGCTGCCCTG 780
QY 261 ProProValLysPheValValGlyAlaIleAlaGlyAlaAsnThrLeuThrGlyIleSer 280
DB 781 CCCCCTGTCAGTGTGTGTAGTGGATTCGAGGCGCAACACCTTGACTGGCATTTCC 840
QY 281 CysGlyLeuLeuAlaSerValAspAlaLeuThrPheGlyGlnPheSerGlyTyrGlyAla 300
DB 841 TGTGGCTTCTAGCCCTCAGTGCCTGACCTTTGGTCTCAGTCTCTGAGTACGAGGCC 900
QY 301 ArgTrpGluThrGlyLeuGlyCysArgAlaThrGlyPheLeuAlaValLeuGlySerGlu 320
DB 901 CGCTGGAGACGGGGCTAGGCTGCCGGGCCACTGGCTTCTTGGCAGTACTTGGGTGGAG 960
QY 321 AlaSerValLeuLeuLeuThrLeuAlaAlaValGlnCysSerValSerValSerCysVal 340
DB 961 GCATCGGTGTGTGTCTCCTCCTGCGCCAGTGCAGTGCAGCGTCTCCGTCTCTGTGTCTC 1020
QY 341 ArgAlaTyrGlyLysSerProSerLeuGlySerValArgAlaGlyValLeuGlyCysLeu 360
DB 1021 CGGGCTATGGGAAGTCCCTCTCCCTGGGCGAGCGTTCGAGCAGGGGTCTTAGGCTGCCGTG 1080
QY 361 AlaLeuAlaGlyLeuAlaAlaLeuProLeuAlaSerValGlyGlyTyrGlyAlaSer 380
DB 1081 GCATCGGAGGGCTGGCGCGCGCTGCTCCCTGGCTCAGTGGGAGATACGGGGCTTCC 1140
QY 381 ProLeuCysLeuProTyrAlaProProGluGlyGlnProAlaAlaLeuGlyPheThrVal 400
DB 1141 CCACTCTGCTGCTTACCGCCACCTGAGGGTTCAGCCAGCAGCGCTGGGTCTTACCGCTG 1200
QY 401 AlaLeuValMetMetAsnSerPheCysPheLeuValValAlaGlyAlaTyrIleLysLeu 420
DB 1201 GCCCTGGTGTATGAATCTCTTCTGTTCCTGGTGGCGGTGGCTTATCAAACTG 1260
QY 421 TyrCysAspLeuProArgGlyAspPheGluAlaValTrpAspCysAlaMetValArgHis 440
DB 1261 TACTGTGACCTTGGCGGGGCGACTTTGAGGCGCTGTGGGACTGGGCCATGGTGGGGC 1320
QY 441 ValAlaTrpLeuIlePheAlaAspGlyLeuLeuTyrCysProValAlaPheLeuSerPhe 460
DB 1321 GTGGCTGGCTCATCTTCGACAGCGGGTCTCTCTACTGTCTCCGTGGCTTCTCAGCTTC 1380
QY 461 AlaSerMetLeuGlyLeuPheProValThrProGluAlaValLysSerValLeuVal 480
DB 1381 GCCTCCATGTGGGCTTCTTCCCTGTACGCGCCGAGCGGCGCTCAAGTCTGTCTGTGTG 1440
QY 481 ValLeuProLeuProAlaCysLeuAsnProLeuLeuTyrLeuLeuPheAsnProHisPhe 500
DB 1441 GTGCTGCCCTTGGCTTCTCAACCCACTGCTGTACCTGCTCTTCAACCCCTTCTTCA 1500
QY 501 ArgAspLeuArgArgLeuArgProArgAlaGlyAspSerGlyProLeuAlaTyrAla 520

Db 1501 CGGATGACCTTCGGCGCTTCGGCCCGCGCAGGGAGCTCAGGCGCCCTAGCCTATGCT 1560
 Qy 521 AlalaGlyGluLeuGluGlySerSerCysAspSerThrGlnAlaLeuValAlaPheSer 540
 Db 1561 CGGCGCGGGAGCTGGAGAGAGCTCCTGTGATTTACCCAGGCGCTGGTAGCCCTTCTCT 1620
 Qy 541 AspValAspLeuLeuLeuAlaSerGluAlaGlyArgProProGlyLeuGluThrTyr 560
 Db 1621 GATGTGGATCTATTCTGGAAGCTTCTGAAGCTGGCGGCGCCCTGGGCTGGAGACCTAT 1680
 Qy 561 GlyPheProSerValThrLeuIleSerCysGlnGlnProGlyAlaProArgLeuGluGly 580
 Db 1681 GGCITCCCTCAGTGACCTCATCTCTGTGTGAGCAGCCAGGCGGCCCGCCAGGCTGGAGGGC 1740
 Qy 581 SerHisCysValGluProGluGlyGlyAsnHisPheGlyAsnProGlnProSerMetAspGly 600
 Db 1741 AGCCATTGTGTAGAGCCAGAGGGAACACATTGGGAACCCCAACCCCTCATGTGATGA 1800
 Qy 601 GlnLeuLeuLeuArgAlaGluGlySerThrProAlaGlyGlyGlyLeuSerGlyGlyGly 620
 Db 1801 GAATCTGCTGTAGGCGCAGAGGGATCTACGCCAGAGGTGGAGGCTTGTGAGGGGGTGGC 1860
 Qy 621 GlyPheGlnProSerGlyLeuAlaPheAlaSerHisVal 633
 Db 1861 GGCITTCAGCCCTCTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCT 1899
 RESULT 4
 AAI67922
 ID AAI67922 standard; cDNA; 2486 BP.
 XX
 AC AAI67922;
 XX
 DT 13-MAR-2002 (first entry)
 XX
 DE Human LGR6 polypeptide encoding cDNA (clone fahr).
 XX
 KW Large G-protein coupled receptor 6; LGR6; GPCR; immunomodulator;
 KW anti-HIV; cycostatic; nootropic; neuroprotective; antiparkinsonian;
 KW anticonvulsant; neuroleptic; antidepressant; tranquilizer; antimigraine;
 KW osteopathic; antiarteriosclerotic; vasotropic; hypotensive; cardiatic;
 KW antiarrhythmic; anorectic; gene therapy; human; ss.
 XX
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FT CDS 2..1903
 FT /*tag= a
 FT /product= "LGR6 polypeptide"
 FT /transl_except= (pos: 170..172, aa: Xaa)
 FT /note= "Xaa = unknown"
 XX
 WO200185768-A2.
 XX
 PN 15-NOV-2001.
 PD
 XX
 PF 08-MAY-2001; 2001WO-US015002.
 XX
 PR 08-MAY-2000; 2000US-00566598.
 XX
 PA (MILL-) MILLENNIUM PHARM INC.
 XX
 PI Gu W;
 XX
 DR WPI; 2002-055584/07.
 DR P-PSDB; AAG66139.
 XX
 PT Novel isolated large G-protein coupled receptor 6 polypeptide, useful for
 PT diagnosing and treating weight disorder, metabolic disorder, central
 PT nervous system disorder, endocrine disorder and cardiovascular disorder.
 XX
 PS Example 1; Fig 4; 198pp; English.
 XX
 XX The invention provides isolated large G-protein coupled receptor 6 (LGR6)

CC polypeptides and polynucleotides encoding them. The LGR6 polypeptides and
 CC polynucleotides are useful as targets for developing modulating agents
 CC that regulate a variety of cellular processes, e.g. neural and endocrine
 CC processes, as well as thermogenesis. They are useful for developing novel
 CC diagnostic and therapeutic agents for LGR6 associated disorders such as
 CC weight disorders (anorexia, obesity), eating disorders, acquired
 CC immunodeficiency syndrome (AIDS)-related wasting, cancer-related wasting,
 CC metabolic disorders, central nervous system disorders (e.g. Alzheimer's
 CC disease, Parkinson's disease, Huntington's disease, Gilles de la
 CC Tourette's syndrome, multiple sclerosis, amyotrophic lateral sclerosis,
 CC epilepsy, Jakob-Creutzfeldt disease, depression, schizoprenia,
 CC obsessive-compulsive disorder, anxiety, panic disorder, migraine),
 CC endocrine disorders (hypo and hyperthyroidism, dwarfism, gigantism,
 CC acromegaly), and cardiovascular disorders (atherosclerosis, reestenosis,
 CC hypertension, myocardial infarction, arrhythmia). The LGR6 polypeptides,
 CC polynucleotides and antibodies are useful in screening assays, detection
 CC assays (e.g., chromosomal mapping, tissue typing, forensic biology),
 CC predictive medicine (e.g. diagnostic assays, prognostic assays,
 CC monitoring clinical trials and pharmacogenomics), and in methods of
 CC treatment (e.g. therapeutic and prophylactic). The present sequence
 CC represents a human LGR6 polypeptide encoding cDNA
 XX

SQ Sequence 2486 BP; 455 A; 773 C; 692 G; 565 T; 0 U; 1 Other;

Alignment Scores:
 Pred. No.: 1,21e-258 Length: 2486
 Score: 3305.00 Matches: 632
 Percent Similarity: 99.84% Conservative: 0
 Best Local Similarity: 99.84% Mismatches: 1
 Query Match: 99.94% Indels: 0
 DB: 6 Gaps: 0

US-10-664-667-5 (1-633) x AAI67922 (1-2486)

Qy 1 AsnThrThrHisTyrArgGluSerTyrTyrAlaCysArgTyrArgSerGlyLeuProGly 20
 Db 2 AATACGACTCATTATAGGAAAGCTGTGTACGCTGCGAGGTACCGTCCGGAATTCCTCGG 61
 Qy 21 SerThrHisAlaSerValGluArgSerGlnGlyLeuSerLeuProAlaHisProAlaSer 40
 Db 62 TCGACCCACCGCTCCGTGGAGCGAGCCAGGGTCTGAGCCTGCGGGCTCATCCAGCCTCT 121
 Qy 41 LeuAlaAlaLeuAlaAlaSerAsnThrThrAlaSerGlyLysLeuGlu***AspThrPhe 60
 Db 122 CTGTGCTCCCTAGCGGCTCCAAACACACCGCATCTGGGAATTTGGAGCTNGACACCTTC 181
 Qy 61 SerGlnLeuSerSerLeuGlnAlaLeuAspLeuSerTrpAsnAlaIleArgSerIleHis 80
 Db 182 AGCCAGCTGAGCTCCCTGCAAGCCCTGGATCTTAGCTGGAAACGCCATCCGGTCCATCCAC 241
 Qy 81 ProGluAlaPheSerThrLeuHisSerLeuValLysLeuAspLeuThrAspAsnGluLeu 100
 Db 242 CCTGAGGCTTCTCCACCCCTGCCTCCCTGGTCAAGCTGAGCTGACAGACCAACACCTG 301
 Qy 101 ThrThrLeuProLeuAlaGlyLeuGlyLeuMetHisLeuLysLeuLysGlyAsnLeu 120
 Db 302 ACCACACTGCCCTGGCTGGACTTGGGGCTTTGATGCATCTGAGCTCAAGGAAACCTT 361
 Qy 121 AlaLeuSerGlnAlaPheSerLysAspSerPheProLysLeuArgIleLeuGluValPro 140
 Db 362 GCTCTCTCCAGGCTTCTCCAAAGGACAGTTTCCCAAACTGAGGATCTCTGGAGTGCCT 421
 Qy 141 TyrAlaTyrGlnCysCysProTyrGlyMetCysAlaSerPhePheLysAlaSerGlyGln 160
 Db 422 TATGCTTACCAGTGTCTCCCTATGGATGTGTGCCAGCTTCTTCAAGGCTCTTGGGCAG 481
 Qy 161 TrpGluAlaGluAspLeuHisLeuAspAspGluGluSerSerLysArgProLeuGlyLeu 180
 Db 482 TGGAGGCTGAAGACCTTACCTTGTATGATGAGAGTCTTCAAAAAGGCCCTTGGGCTC 541
 Qy 181 LeuAlaArgGlnAlaGluAsnHisTyrAspGlnAspLeuAspGluLeuGluMet 200
 Db 542 CTTGCCAGACACAGAGAACCACTATGACAGGACCTGAGCTGAGTCCAGTCCAGTGGAGATG 601

QY 201 GluAspSerLysProHisProSerValGlnCysSerProThrProGlyProPhelysPro 220
DB 602 GAGGACTCAAGCCACACCCAGTGTCCAGTGTAGCCCTACTCCAGGCCCTTCAAGCCC 661
QY 221 CysGluTyrLeuPheGluSerTrpGlyIleArgLeuAlaValTrpAlaIleValLeuLeu 240
DB 662 TGAGTAGTACCTCTTTGAAAGCTGGGGCATCCGCCCTGGCCGTGGGGCCATCGTGTGCTC 721
QY 241 SerValLeuCysAsnGlyLeuValLeuLeuThrValPheAlaGlyGlyProAlaProLeu 260
DB 722 TCCGTGCTCTGCANATGGACTGGTGTCTGTACCGGTGTTCGCTGGCGGGCTGCCCCCTG 781
QY 261 ProProValLysPheValValGlyAlaIleAlaGlyAlaAsnThrLeuThrGlyIleSer 280
DB 782 CCCCCTGCAAGTTTGTGTAGTGGATTTGCAGGTTCAGCGGCCAACACACTTGACTGGCATTTCC 841
QY 281 CysGlyLeuLeuAlaSerValAspAlaLeuThrPheGlyGlnPheSerGluTyrGlyAla 300
DB 842 TGTGGCCTTTAGCCCTCAGTCAGTCAGCCCTGACCTTTGGTTCAGTTCTCTGAGTACGGAGCC 901
QY 301 ArgTyrGluThrGlyLeuGlyCysArgAlaThrGlyPheLeuAlaValLeuGlySerGlu 320
DB 902 CGCTGGAGACGGGCTAGGCTGGCCGGCCACTGGCTTCTGGCAGTACTTGGGTGGAG 961
QY 321 AlaSerValLeuLeuLeuThrLeuAlaAlaValGlnCysSerValSerCysVal 340
DB 962 GCATCGGTGCTGCTCTCACTCTGCGCCGAGTGCAGTGCAGGGTCTCGGTCTCTGTGTC 1021
QY 341 ArgAlaTyrGlyLysSerProSerLeuGlySerValArgAlaGlyValLeuGlyCysLeu 360
DB 1022 CGGGCTATGGAGTCCGCTCCCTCGGCGAGCGTTCGAGCAGGGGTCTAGGCTGCCTG 1081
QY 361 AlaLeuAlaGlyLeuAlaAlaLeuProLeuAlaSerValGlyGlyTyrGlyAlaSer 380
DB 1082 GCACTGGCAGGGCTGGCCCGCCACTGCCCTGGCCCTCAGTGGGAGAAATACGGGCCCTCC 1141
QY 381 ProLeuCysLeuProTyrAlaProGluGlyGlnProAlaAlaLeuGlyPheThrVal 400
DB 1142 CCATCTGCTGCTGCTAGCGGCCACTTGAGGGTACAGCCAGCAGCCCTGGGCTTCCACCGTG 1201
QY 401 AlaLeuValMetMetAsnSerPheCysPheLeuValValAlaGlyAlaTyrIleLysLeu 420
DB 1202 GCCCTGGTGCATGATGAATCTCTCTGTTTCTGGTCTGGCGGTGCCTATACAAACTG 1261
QY 421 TyrCysAspLeuProArgGlyAspPheGluAlaValTrpAspCysAlaMetValArgHis 440
DB 1262 TACTGTGACCTGCGCGGGGCGACTTTTGGGCCGTGTGGACTGCGCCATGGTGGGCAC 1321
QY 441 ValAlaTrpLeuIlePheAlaAspGlyLeuLeuTyrCysProValAlaAlaPheLeuSerPhe 460
DB 1322 GTGGCTGGCTCATCTTGCAGACGGGCTCTCTACTGTCTCCGTGGCTTCTCTAGCTTC 1381
QY 461 AlaSerMetLeuGlyLeuPheProValThrProGluAlaValLysSerValLeuVal 480
DB 1382 GCCTCCATGCTGGGCTCTTCCCTGTACGCCCGAGGCCGTCAAGTCTGTCTGTGCTGTG 1441
QY 481 ValLeuProLeuProAlaCysLeuAsnProLeuLeuTyrLeuLeuPheAsnProHisPhe 500
DB 1442 GTGCTGCCCTGCTGCTGCTCAACCCACTGCTGTACCTGCTCTTCAACCCCCACTTC 1501
QY 501 ArgAspLeuArgLeuArgProArgAlaGlyAspSerGlyProLeuAlaTyrAla 520
DB 1502 CGGGATGACCTTGGCGGCTTCCGGCCCGCGCAGGGGACTCAGGGCCCTTAGCTATGCT 1561
QY 521 AlaAlaGlyLeuGlyLysSerSerCysAspSerThrGlnAlaLeuValAlaPheSer 540
DB 1562 GCGCGCGGGAGCTGGAGAGAGCTCTCTGTGATTCTACCCAGGCCCTGTGTAGCTTCTCT 1621
QY 541 AspValAspLeuIleLeuGluAlaSerGluAlaGlyArgProProGlyLeuGluThrTyr 560
DB 1622 GATGTGGATCTCATTTCTGAAAGCTTTCTGAAGCTGGGGCGGCCCTGGGCTGGAGACCTAT 1681

QY 561 GlyPheProSerValThrLeuIleSerCysGlnGlnProGlyAlaProArgLeuGluGly 580
DB 1682 GGCCTCCCTCAGTGACCTCATCTCTGTCTCAGCAGCCAGGGGCCCCAGGCTGGAGGC 1741
QY 581 SerHisCysValGluProGluGlyAsnHisPheGlyAsnProGlnProSerMetAspGly 600
DB 1742 AGCCATTGTGTAGAGCCAGAGGGGAACACACTTTGGGAACCCCAACCTCTCCATGGATGA 1801
QY 601 GluLeuLeuLeuArgAlaGluGlySerThrProAlaGlyGlyGlyLeuSerGlyGly 620
DB 1802 GAATGCTGCTGAGGGCAGAGGGATCTACGCCAGCAGGTGGAGGCTTGTCTCAGGGGTGGC 1861
QY 621 GlyPheGlnProSerGlyLeuAlaPheAlaSerHisVal 633
DB 1862 GGCCTTTCAGCCCTCTGGCTTGGCTTGGCTTTCACACGCTG 1900
RESULT 5
ADN02242
ID ADN02242 standard; cDNA; 2486 BP.
XX
AC ADN02242;
XX
DT 17-JUN-2004 (first entry)
XX
DE Human partial large G protein-coupled receptor, LGR6, cDNA #1.
XX
KW Human; ss; large G protein-coupled receptor 6; LGR6; heart failure;
KW ischaemic heart disease; myocardial infarction; hypertension;
KW pericarditis; atherosclerosis; hepatic failure; viral hepatitis; cancer;
KW intracerebral haemorrhage; brain abscess; Alzheimer's disease;
KW Pick disease; Parkinson's disease; Huntington's disease;
KW amyotrophic lateral sclerosis; inflammatory bowel disease;
KW Crohn's disease; benign prostatic hypertrophy;
KW systemic lupus erythematosus.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 2..1903
FT /tag= a
FT /partial
FT /notes= "No start codon shown"
FT /product= "partial LGR6"
XX
PN US2003166047-A1.
XX
PD 04-SEP-2003.
XX
PF 08-MAY-2001; 2001US-00851595.
XX
PR 06-MAY-1999; 99US-0132896P.
PR 08-MAY-2000; 2000US-00566588.
XX
PA (MILL-) MILLENNIUM PHARM INC.
XX
PI Gu W;
XX
DR WPI; 2003-898067/82.
DR P-PSDB; ADN02243.
XX
PT New isolated LGR6 nucleic acids and polypeptides, useful for diagnosing
PT and treating LGR6-mediated disorders, such as myocardial infarction,
PT hypertension, atherosclerosis, viral hepatitis, cancer and/or Alzheimer's
PT disease.
XX
PS Claim 1; SEQ ID NO 4; 145pp; English.
XX
CC The invention relates to an isolated LGR6 (large G protein-coupled
CC receptor) nucleic acid molecule comprising the mouse and human cDNA
CC sequences appearing as ADN02239, ADN02241, ADN02242, ADN02244, ADN02245,
CC ADN02247, ADN02248 and ADN02250. Also included are a host cell or a non-
CC human mammalian host cell containing an LGR6 nucleic acid (and producing
CC LGR6), an isolated LGR6 polypeptide appearing as ADN02240, ADN02243,

Db 1742 AGCCATTGTGTAGCCAGAGGGGAACCACTTTGGGAACCCCAACCTCCATGGATGGA 1801
Qy 601 GluLeuLeuLeuAlaGlySerThrProAlaGlyGlyGlyLeuSerGlyGlyGly 620
Db 1802 GAACCTGCTGTAGGGCAGAGGATCTAGCCAGCAGGTGGAGGCTTGTTCAGGGGGTGGC 1861
Qy 621 GlyPheGlnProSerGlyLeuAlaPheAlaSerHisVal 633
Db 1862 GGCTTTACGCCCTGTGGCTTGGCCCTTGTTCACACGTG 1900

RESULT 6

ADK19407

ID ADK19407 standard; cDNA; 2486 BP.

XX ADK19407;

DT 17-JUN-2004 (first entry)

XX cDNA encoding human LGR6 polypeptide #1.

DE G-protein coupled receptor; GPCR; large G-protein coupled receptor 6;
XX LGR6; weight disorder; neural disorder; endocrine disorder;
KW cardiovascular disorder; anorectic; anabolic; immunomodulator;
KW neuroprotective; nootropic; antiparkinsonian; anti-HIV; anticonvulsant;
KW antithyroid; osteopathic; cardiant; vasotropic; hypotensive;
KW arteriosclerotic; antiarrhythmic; human; gene; ss.

XX Homo sapiens.

XX Key Location/Qualifiers

FT CDS 2..1903

FT /tag= a

FT /partial

FT /product= "LGR6 polypeptide #1"

FT /note= "This sequence lacks a start codon"

FT /transl_except= (pos:170..172,aa:Xaa)

FT /note= "Xaa= Unknown"

XX US2004058377-A1.

XX 25-MAR-2004.

XX 18-SEP-2003; 2003US-00664667.

XX 06-MAY-1999; 99US-0132896P.

XX 08-MAY-2000; 2000US-00566588.

XX 08-MAY-2001; 2001US-00851595.

XX (MILL-) MILLENNIUM PHARM INC.

XX Gu W;

XX WPI; 2004-268789/25.

XX P-PSDB; ADK19408.

XX New large G-protein coupled receptor 6 nucleic acid molecules and
PT polypeptides useful for diagnosing, preventing or treating diseases
PT associated with aberrant nucleic acid or protein activity, e.g. obesity,
PT anorexia or cachexia.

XX Claim 2; SEQ ID NO 4; 145pp; English.

XX The present invention relates to the isolation of novel members of the G-
CC protein coupled receptor (GPCR) family designated as large G-protein
CC coupled receptor 6 (LGR6), and the polynucleotide sequences encoding
CC them. The invention also discloses LGR6 fusion proteins, antigenic
CC peptides, anti-LGR6 antibodies, recombinant expression vectors, host
CC cells, and non-human transgenic animals in which an LGR6 gene has been
CC introduced or disrupted. The LGR6 polynucleotide sequences, polypeptide
CC sequences, and antibodies are useful for diagnosing, treating or
CC preventing diseases associated with aberrant LGR6 expression or activity,
CC such as weight disorders (e.g. obesity, anorexia nervosa or cachexia),

CC neural disorders (e.g. central nervous system disorder, including
CC Alzheimer's disease, Parkinson's disease, AIDS-related dementia, multiple
CC sclerosis and Huntington's disease), endocrine disorders (e.g.
CC hypothyroidism, hyperthyroidism or acromegaly), or cardiovascular
CC disorders (e.g. ischaemia-reperfusion injury, coronary artery disease,
CC myocardial infarction, arrhythmia, atherosclerosis, hypertension or
CC congestive heart failure). The LGR6 polynucleotide and polypeptide
CC sequences are also useful as targets for developing modulating agents
CC that regulate a variety of cellular processes, e.g. neural and endocrine
CC processes and thermogenesis. They can be used in screening assays (e.g.
CC chromosome mapping, tissue typing or in forensic biology), or in
CC predictive medicine (e.g. diagnostic assays, prognostic assays,
CC monitoring clinical trials or pharmacogenomics). The present sequence
CC encodes a human LGR6 polypeptide.

SQ Sequence 2486 BP; 455 A; 773 C; 692 G; 565 T; 0 U; 1 Other;

Alignment Scores:

Pred. No.: 1.21e-258 Length: 2486
Score: 3305.00 Matches: 632
Percent Similarity: 99.84% Conservative: 0
Best Local Similarity: 99.84% Mismatches: 1
Query Match: 99.94% Indels: 0
DB: 12 Gaps: 0

US-10-664-667-5 (1-633) x ADK19407 (1-2486)

Qy 1 AsnThrThrHisTyrArgGluSerTrpTyrAlaCysArgTyrArgSerGlyIleProGly 20
Db 2 AATACGACTCTACTATAGGGAAGCTGGTACGCCCTGCAGTACCGGTCCGGAATTCCTCCGGG 61
Qy 21 SerThrHisAlaSerValGluArgSerGlnGlyLeuSerLeuProAlaHisProAlaSer 40
Db 62 TCGACCCACGCGTCCGTGGAGCGGAGCCAGGGTCTGAGCCTGCCGCTCATCCAGCCTCT 121
Qy 41 LeuAlaLeuAlaAlaSerAsnThrThrAlaSerGlyLysLeuGlu***AspThrPhe 60
Db 122 CTGTGTCCTTAGCGGCTCCCAACACACACCGCATCTGGGAAATTGGAGCTNGACACCTTC 181
Qy 61 SerGlnLeuSerSerLeuGlnAlaLeuAspLeuSerTrpAsnAlaIleArgSerIleHis 80
Db 182 AGCCAGCTGAGCTCCCTGCAAGCCCTGGATCTTAGCTGGAACGCCATCCGGTTCATCCAC 241
Qy 81 ProGluAlaPheSerThrLeuHisSerLeuValLysLeuAspLeuThrAspAsnGlnLeu 100
Db 242 CCTGAGGCTTCTCCACCCCTGCACTCCCTGCTCAAGCTGGAGCTGACAGACACACCGCTG 301
Qy 101 ThrThrLeuProLeuAlaGlyLeuGlyGlyLeuMetHisLeuLysLeuLysGlyAsnLeu 120
Db 302 ACCACACTGCCCTGGCTGGACTTGGGGGCTTGATGCATCTGAAAGCTCAAAGGGAACCTT 361
Qy 121 AlaLeuSerGlnAlaPheSerLysAspSerPheProLysLeuArgIleLeuGluValPro 140
Db 362 GCTCTCTCCAGGCTTCTCCCAAGACAGATTTCCCAAACTGAGGATCCTGGAGGTGCT 421
Qy 141 TyrAlaTyrGlnCysCysProTyrGlyMetCysAlaSerPhePheLysAlaSerGlyGln 160
Db 422 TATGCTACCACTGCTGCTCCCTATGGGATGTGTCCAGCTTCTTCAAGGCTCTGGGAG 481
Qy 161 TrpGluAlaGluAspLeuHisLeuAspAspGluGluSerSerLysArgProLeuGlyLeu 180
Db 482 TGGGAGGCTGAAGACCTTACCTTGTATGATGAGGAGTCTTCAAAAGGCCCTCGGCCCTC 541
Qy 181 LeuAlaArgGlnAlaGluAsnHisTyrAspGlnAspLeuAspGluLeuGluMet 200
Db 542 CTTGCCAGACAAAGCAGAGAACCACTATGACAGACCTGGATGAGCTCCAGCTGGAGATG 601
Qy 201 GluAspSerLysProHisProSerValGlnCysSerProThrProGlyProPheLysPro 220
Db 602 GAGGACTCAAGGCCACACCCCGCTGTCAGTGTAGCCCTACTCCAGGCGCTTCAAGGCC 661
Qy 221 CysGluTyrLeuPheGluSerTrpGlyIleArgLeuAlaValTrpAlaIleValLeuLeu 240

Db 662 TGTGAGTACCTCTTTAAAGCTGGGGCATCCGCTGGCCGCTGTGGCCATCGTGTGTGCTC 721
Qy 241 SerValLeuCysAsnGlyLeuValLeuLeuThrValPheAlaGlyGlyProAlaProLeu 260
Db 722 TCCGTCCTCTGCAATGGACTGGTGTGTGTCACCGTGTTCGTGGCGGCCCTGCCCCCTGTG 781
Qy 261 ProProValLysPheValValGlyAlaIleAlaGlyAlaAsnThrLeuThrGlyIleSer 280
Db 782 CCCCCTGTCAGATTTGTGTAGTGTGAGTGTGAGGCGCCAAACACCTTGACTGGCAATTC 841
Qy 281 CysGlyLeuLeuAlaSerValAspAlaLeuThrPheGlyGlnPheSerGluTyrGlyAla 300
Db 842 TGTGGCTTCTAGCCTCAGTCAGTCAGCTGACCTTTTGTTCAGTTCTCTGAGTACGGAGCC 901
Qy 301 ArgTrpGluThrGlyLeuGlyCysArgAlaThrGlyPheLeuAlaValLeuGlySerGlu 320
Db 902 CGCTGGGAGACGGGGCTAGGCTGCGGGCCACTGGCTTCTCTGGCAGTACTTGGGTGGAG 961
Qy 321 AlaSerValLeuLeuLeuThrLeuAlaAlaValGlnCysSerValSerValSerCysVal 340
Db 962 GCATCGGTGCTGTGCTCATCTCTGGCCGACGTGACGTGCAGCGTCTCCGTCTCTGTGTC 1021
Qy 341 ArgAlaTyrGlyLysSerProSerLeuGlySerValArgAlaGlyValLeuGlyCysLeu 360
Db 1022 CGGGCTATGGGAAGTCCCTCCCTGGCGAGCGTTGAGCAGGGGTCTTAGGCTGCCTG 1081
Qy 361 AlaLeuAlaGlyLeuAlaAlaLeuProGluLeuAlaSerValGlyGluTyrGlyAlaSer 380
Db 1082 GCACCTGGCAGGGCTGGCGCGGCACCTGCCCTCGCTCAGTGGGAGAAATACGGGGCTCC 1141
Qy 381 ProLeuCysLeuProTyrAlaProProGluGlyGlnProAlaAlaLeuGlyPheThrVal 400
Db 1142 CCATCTCCCTGCTCCCTACCGCCACCTGAGGGTCAGCCAGCAGCCCTGGGCTTACCGTG 1201
Qy 401 AlaLeuValMetMetAsnSerPheCysPheLeuValAlaAlaGlyAlaTyrIleLysLeu 420
Db 1202 GCCCTGGTGATGATGAATCTCTCTGTTTCTGGTGGCGGTGCGCTACATCAAACTG 1261
Qy 421 TyrCysAspLeuProArgGlyAspPheGluAlaValTrpAspCysAlaMetValArgHis 440
Db 1262 TACTGTGACCTGCCCGGGGGACCTTTGAGGCCGTGTGGGACTGCGCCATGGTGGGCAC 1321
Qy 441 ValAlaTrpLeuIlePheAlaAspGlyLeuLeuTyrCysProValAlaPheLeuSerPhe 460
Db 1322 GTGGCTGGCTCACTCTTCGACAGCGGGCTCTCTACTGTCCGTGGCTTCTCTCAGCTTC 1381
Qy 461 AlaSerMetLeuGlyLeuPheProValThrProGluAlaValLysSerValLeuVal 480
Db 1382 GCCTCCATGCTGGGCTCTCTCCCTGTACGCCCGCGGCGTCAAGTCTGTCTGCTGGTG 1441
Qy 481 ValLeuProLeuProAlaCysLeuAsnProLeuLeuTyrLeuLeuPheAsnProHisPhe 500
Db 1442 GTGCTGCCCTCGCTCGCTGCTCAACCCACTGCTGTACCTGCTCTTCAACCCCCACTTC 1501
Qy 501 ArgAspLeuLeuArgArgLeuArgProArgAlaGlyAspSerGlyProLeuAlaTyrAla 520
Db 1502 CGGATGACCTTCGGCGGCTTCGGCCCCGCGCAGGGACTCAGGGCCCTAGCCTATGCT 1561
Qy 521 AlaAlaGlyGluLeuGluLysSerSerCysAspSerThrGlnAlaLeuValAlaPheSer 540
Db 1562 CGCGCGGGGAGCTGGAGAAGAGTCTCTGTGATTCACCCAGGCCCTGGTAGCTTCTCT 1621
Qy 541 AspValAspLeuIleLeuGluAlaSerGluAlaGlyArgProProGlyLeuGluThrTyr 560
Db 1622 GATGTGATCTCATTTCTGAAAGCTTCTGAAGCTGGCGGCCCTCGCTGGGACCTAT 1681
Qy 561 GlyPheProSerValThrLeuIleSerCysGlnGlnProGlyAlaProArgLeuGluGly 580
Db 1682 GCCTTCCCTCAGTGACCTCATCTCTCTGTGACAGCCAGGGGCCCCAGGCTGGAGGC 1741
Qy 581 SerHisCysValGluProGluGlyAsnHisPheGlyAsnProGlnProSerMetAspGly 600
Db 1742 AGCCATTGTGTAGAGCCAGAGGGGAACCACTTTGGGAACCCCAACCCCTCCATGGATGA 1801

Qy 601 GluLeuLeuLeuArgAlaGluGlySerThrProAlaGlyGlyLeuSerGlyGlyGly 620
Db 1802 GAATGCTGCTGAGGCGCAGAGGATCTACCCAGCAGTGGAGCTTGTACGGGGTGGC 1861
Qy 621 GlyPheGlnProSerGlyLeuAlaPheAlaSerHisVal 633
Db 1862 GGCTTCAGGCCCTGCGCTTGGCTTGGCTTTCACACGTG 1900

RESULT 7

AAF98722
ID AAF98722 standard; DNA; 2322 BP.
AC AAF98722;
XX 02-JUL-2001 (first entry)
DT Human late stage ovarian tumour polynucleotide marker 37.
DE Human; ovarian cancer; identification; detection; characterisation;
KW tumour; kinase; marker; cytostatic; antisense gene therapy; ds.
XX Homo sapiens.

OS
XX WO200118542-A2.
FN
XX 15-MAR-2001.
PD
XX 01-SEP-2000; 2000WO-US024199.
PF
XX 03-SEP-1999; 99US-0152547P.
PR 16-MAR-2000; 2000US-0190347P.
PR 21-MAR-2000; 2000US-0191321P.
PR 31-MAY-2000; 2000US-0208382P.
PR 20-JUL-2000; 2000US-00220467.
XX
XX (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.

PI Lee J, Thompson P, Lillie J;
XX WPI; 2001-211428/21.
DR
XX

PT Detection, assessment, prevention and therapy of ovarian cancer,
PS comprises detecting changes in the expression of a variety of markers.

XX Claim 1; Page 1188-1189; 1198pp; English.

XX The present invention describes a method for assessing whether a patient is afflicted with ovarian cancer by comparing: (1) the expression of a marker (1) (see AAF98594 to AAF98730), in a patient sample; and (2) the normal level of expression of (1) in a control non-ovarian cancer sample, where a significant difference between the level of expression in (a) and (b) is an indication that the patient is afflicted with ovarian cancer. (1) have cytostatic activities and can be used in antisense gene therapy. The method, compositions and kits from the present invention can be used for: (1) assessing and treating ovarian cancer; (2) making isolated hybridoma, which produces an antibody useful for ovarian cancer assessment; and (3) inhibiting ovarian cancer in a patient. AAF98573 to AAF98593 represent human kinase marker primers and probes which are used in the exemplification of the present invention

XX Sequence 2322 BP; 407 A; 733 C; 652 G; 530 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.: 4,46e-246 Length: 2322
Score: 3150.00 Matches: 609
Percent Similarity: 99.84% Conservative: 1
Best Local Similarity: 99.67% Mismatches: 1
Query Match: 95.25% Indels: 1
DB: 5 Gaps: 0

US-10-664-667-5 (1-633) x AAF98722 (1-2322)

QY 23 HisAlaSerValGluArgSerGlnGlyLeuSerLeuProAlaHisProAlaSerLeuAla 42
DB 2 CACGGCTCCGTCAGCGGAGCAGAGGTCCTGAGCCTCGCGGCTCATCCAGCCTCTCTTGCT 61
QY 43 AlaLeuAlaAlaSerAsnThrThrAlaSerGlyLysLeuGlu**AspThrPheSerGln 62
DB 62 GCCCTAGCGGCTCCAAACACCGCATCTGGGAATTTGAGCT-GACACCTTCAGCCAG 120
QY 63 LeuSerSerLeuGlnAlaLeuAspLeuSerTrpAsnAlaIleArgSerIleHisProGlu 82
DB 121 CTGAGCTCCCTCAAGCCCTGATCTTAGCTGGAACGCCATCCGGTCCATCCACCTGAG 180
QY 83 AlaPheSerThrLeuHisSerLeuValLysLeuAspLeuThrAspAsnGlnLeuThrThr 102
DB 181 GCCTTCTCCACCCCTGCACCTCCCTGGTCAAGCTGGACCTGACAGACAACAGCAGTGCACCA 240
QY 103 LeuProLeuAlaGlyLeuGlyLeuMetHisLeuLysLeuLysLeuLysLeuAlaLeu 122
DB 241 CTGCCCTGGCTGGACTTGGGGCTTGATGCATCTGAAGCTCAAGGGNAACCTTGTCTTC 300
QY 123 SerGlnAlaPheSerLysAspSerPheProLysLeuArgIleLeuGluValProTyrAla 142
DB 301 TCCAGGCCCTTCCAAAGACAGTTCCTCCAAACTGAGGATCCTGAAGGTGCTTATGCC 360
QY 143 TyrGlnCysCysProTyrGlyMetCysAlaSerPhePheLysAlaSerGlyGlnTrpGlu 162
DB 361 TACCAGTGTGTCCTATGGGATGTGTCCAGCTTCTTCAAGGCCCTCTGGGCAGTGGGAG 420
QY 163 AlaGluAspLeuHisLeuAspLeuGluSerSerLysArgProLeuGlyLeuLeuAla 182
DB 421 GCTGAAGACCTTCACCTTGATGAGGAGTCTTCAAAAGGCCCTGGGGCTCTCTTGCC 480
QY 183 ArgGlnAlaGluAsnHisTyrAspGlnAspLeuAspGluLeuGlnLeuMetGluAsp 202
DB 481 AGACAGCAGAGAACCATATACACAGGACCTGGATGAGCTCCAGCTGGAGATGGAGGAC 540
QY 203 SerLysProHisProSerValGlnCysSerProThrProGlyProPheLysProCysGlu 222
DB 541 TCAAGGCCACACCCAGTGTCCAGTGTAGCCCTACTCCAGGCCCTTCAAGCCCTGTGAG 600
QY 223 TyrLeuPheGluSerTrpGlyIleArgLeuAlaValTrpAlaIleValLeuLeuSerVal 242
DB 601 TACCTCTTTGAAAGCTGGGGCATCCGCTGGCGGTGGGCCCATCGTGTGCTCTCCGTG 660
QY 243 LeuCysAsnGlyLeuValLeuLeuThrValPheAlaGlyGlyProAlaProLeuProPro 262
DB 661 CTCTGCATGGACTGGTGTGCTGACCGTTCGCTGGCGGCCCTGCCCTCCCTGCCCGG 720
QY 263 ValLysPheValValGlyAlaIleAlaGlyAlaAsnThrLeuThrGlyIleSerCysGly 282
DB 721 GTCAAGTTTGTAGGTGCGATTGACGGCGCCACACACCTTGACCTGGCATTTCTGTGGC 780
QY 283 LeuLeuAlaSerValAspAlaLeuThrPheGlyGlnPheSerGlyTyrGlyAlaArgTrp 302
DB 781 CTCTAGCCTCAGTCGATGCCCTGACCTTTGTGTCAGTTCTCTGAGTACGGAGCCCGCTGG 840
QY 303 GluThrGlyLeuGlyCysArgAlaThrGlyPheLeuAlaValLeuGlySerGluAlaSer 322
DB 841 GAGACGGGCTAGGCTGCGGGCCACTGGCTTCCTGGCAGTACTTGGGTGGAGGATCG 900
QY 323 ValLeuLeuLeuThrLeuAlaAlaValGlnCysSerValSerValSerCysValArgAla 342
DB 901 GTGCTGTGCTCACTCTGGCCGACGTGACGTGACGTGCTCGTCTCTGCTGTCTCCGGCC 960
QY 343 TyrGlyLysSerProSerLeuGlySerValArgAlaGlyValLeuGlyCysLeuAlaLeu 362
DB 961 TATGGGAAGTCCCTCCCTGGGACGCTTCGAGCAGGGGCTCTAGGCTGCCCTGGCACTG 1020
QY 363 AlaGlyLeuAlaAlaLeuProLeuAlaSerValGlyGlyThrGlyAlaSerProLeu 382
DB 1021 GAGGGCTGGCCGCGCACTGCCCTTGGCCTCAGTGGGAGAATACGGGGCCCTCCCACTC 1080
QY 383 CysLeuProTyrAlaProProGluGlyGlnProAlaAlaLeuGlyPheThrValAlaLeu 402

DB 1081 TGCTGCCCTACGGCCACCTGAGGGTCAGGCAGAGCCCTGGGCTTCCCGTGGCCCTG 1140
QY 403 ValMetMetAsnSerPheCysPheLeuValValAlaGlyAlaTyrIleLysLeuTyrCys 422
DB 1141 GTGATGATGAACCTCTTCTGTTCTGTCGTGGCGGTGCTACATCANACTGTACTGT 1200
QY 423 AspLeuProArgGlyAspPheGluAlaValTrpAspCysAlaMetValArgHisValAla 442
DB 1201 GACCTGCCCGGGCGACTTTGAGCCCGTGGGACTGCGCATGCTGAGGACGCTGGCC 1260
QY 443 TrpLeuIlePheAlaAspGlyLeuLeuTyrCysProValAlaPheLeuSerPheAlaSer 462
DB 1261 TGGCTCACTCTCGCAGACGGGCTCCTACTGTCCCGTGGGCTTCTCAGCTTCGCCCTC 1320
QY 463 MetLeuGlyLeuPheProValThrProGluAlaValLysSerValLeuLeuValValLeu 482
DB 1321 ATGTGGGCCCTTCCCTGTCAAGCCCGAGGCCGCTCAAGTCTGTCTGCTGCTGTGCTG 1380
QY 483 ProLeuProAlaCysLeuAsnProLeuLeuTyrLeuLeuPheAsnProHisPheArgAsp 502
DB 1381 CCCCTGCTGCTGCTCAACCCACTGCTGTACCTGCTCTTCAACCCCCACTTCGGGGAT 1440
QY 503 AspLeuArgArgLeuArgProArgAlaGlyAspSerGlyProLeuAlaTyrAlaAlaAla 522
DB 1441 GACCTTCCGGCGCTTCGGCCCGCGCAGGGGACTCAGGGGCCCTAGCCTATGCTCGGCC 1500
QY 523 GlyGluLeuGluLysSerSerCysAspSerThrGlnAlaLeuValAlaPheSerAspVal 542
DB 1501 GGGAGCTGGAGAGAGCTCTCTGATTTACCCAGGCCCTGTTAGCTTCTCTGATGTG 1560
QY 543 AspLeuIleLeuGluAlaSerGluAlaGlyArgProProGlyLeuGluThrTyrGlyPhe 562
DB 1561 GATCTCATCTGGAAGCTTCTGAAGCTGGCGGCCCTTGGGCTGGAGACCTATGGCTTC 1620
QY 563 ProSerValThrLeuIleSerCysGlnGlnProGlyAlaProArgLeuGluGlySerHis 582
DB 1621 CCCTCAGTGCCTCATCTCTGTGTCAGCAGCCAGGGGCCCGCCAGGCTGGAGGGCAGCCAT 1680
QY 583 CysValGluProGluGlyAsnHisPheGlyAsnProGlnProSerMetAspGlyGluLeu 602
DB 1681 TGTGTAGAGCCAGAGGGGAACACATTTGGGAACCCCCAACCCCTCATGGATGGAGACTG 1740
QY 603 LeuLeuArgAlaGluGlySerThrProAlaGlyGlyGlyLeuSerGlyGlyGlyPhe 622
DB 1741 CTGCTGAGGGCAGAGGGATCTACGCCACAGGTGGAGGCTTGTACAGGGGGTGGCGGCTTT 1800
QY 623 GlnProSerGlyLeuAlaPheAlaSerHisVal 633
DB 1801 CAGCCCTCTGGCTGGCCCTTGTCTTCACACGTG 1833

RESULT 8

ADQ76775
ID ADQ76775 standard; DNA; 3451 BP.
XX AC ADQ76775;
XX DT 07-OCT-2004 (first entry)
XX DE Human LGR6 splice variant LGR6.2 nucleotide sequence.
XX KW Human; leucine-rich repeat-containing G-protein coupled receptor;
KW receptor; LGR6.2; cytosolic; immunomodulator; muscular-gen;
KW synaenological; antinfertility; endocrine-gen.; anabolic; hypertensive;
KW vulnerary; dermatological; gene therapy; gene; ds.
XX OS Homo sapiens.
XX FH Key Location/Qualifiers
FT CDS 47..1093
FT FT /*tag= b
FT FT /product= "Human LGR6.2"
FT FT sig_peptide 47..103

QY 407 SerPheCysPheLeuValValAlaGlyAlaTyrIleLeuTyrCysAspLeuProArg 426
 Db 2214 TCCTTCTGTTCCTGCTGCGCGTGCCTACATCAAACTGTACTGTGACCTGCGCGG 2273
 QY 427 GlyAspPheGluAlaValTrpAspCysAlaMetValArgHisValAlaTrpLeuIlePhe 446
 Db 2274 GCGGACTTTGAGCGCTGTGGGACGTGCGCATGTGGTGGACGCTGGCTGCCTCATCTTC 2333
 QY 447 AlaAspGlyLeuLeuTyrCysProValAlaPheLeuSerPheAlaSerMetLeuGlyLeu 466
 Db 2334 CGACAGCGGCTCTCTACTGTCCCGTGGCTTCTCAGCTTCGCTCCATGCTGGGCTC 2393
 QY 467 PheProValThrProGluAlaValLysSerValLeuValValLeuProLeuProAla 486
 Db 2394 TTCCTGTGACGCGCGGCGCTGCAAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2453
 QY 487 CysLeuAsnProLeuLeuTyrLeuLeuPheAsnProHisPheArgAspLeuArg 506
 Db 2454 TGCTTCACCCACTGCTGTACTGCTCTTCAACCCCACTTCGGGATGACCTTCGGCGG 2513
 QY 507 LeuArgProArgAlaGlyAspSerGlyProLeuAlaTyrAlaAlaGlyLeuLeuGlu 526
 Db 2514 CTTGGCCCCGCGGAGGACCTCAGGCCCCCTAGCTATGCTGCGGCGGGAGCTGGAG 2573
 QY 527 LysSerSerCysAspSerThrGlnAlaLeuValAlaPheSerAspValAspLeuLeu 546
 Db 2574 AAGAGCTCTGTGATTTACCCAGGCGCTGTGAGCTTCTCTGATGTGGATCTCATCTG 2633
 QY 547 GluAlaSerGluAlaGlyArgProProGlyLeuGluThrTyrGlyPheProSerValThr 566
 Db 2634 GAAGCTTTCTGAAGCTGGGCGGCCCCCTGGGCTGGAGACCTATGGCTTCCCTCAGTGACC 2693
 QY 567 LeuIleSerCysGlnGlnProGlyAlaProArgLeuGluGlySerHisCysValGluPro 586
 Db 2694 CTCATCTCTGTGACGAGCCAGGCGCCCCAGGCTGGAGGCGAGCCATTGTGTAGAGCA 2753
 QY 587 GluGlyAsnHisPheGlyAsnProGlnProSerMetAspGlyGluLeuLeuArgAla 606
 Db 2754 GAGGGGAACCACTTTGGGAACCCCAACCCCTCATGGATGGAGAACTGTGCTGAGGGCA 2813
 QY 607 GluGlySerThrProAlaGlyGlyLeuSerGlyGlyGlyPheGlnProSerGly 626
 Db 2814 GAGGGATCTACGCCAGCGTGGAGCTTGTGAGGGGTGGCGGCTTTCAGCCCTCTGCG 2873
 QY 627 LeuAlaPheAlaSerHisVal 633
 Db 2874 TTGGCTTTGCTTCACACGTG 2894
 RESULT 9
 ID AAI67925
 XX AAI67925 standard; cDNA; 2208 BP.
 AC AAI67925;
 XX
 DT 13-MAR-2002 (first entry)
 XX
 DE Human LGR6 polypeptide coding sequence.
 XX
 KW Large G-protein coupled receptor 6; LGR6; GPCR; immunomodulator;
 KW anti-HIV; cytostatic; neotropic; neuroprotective; antiparkinsonian;
 KW anticonvulsant; neuroleptic; antidepressant; tranquilizer; antimigraine;
 KW osteopathic; antiarteriosclerotic; vasotropic; hypotensive; cardiant;
 KW antiarrhythmic; anorectic; gene therapy; human; ss.
 XX
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FT CDS 1..2208
 FT /*tag= a
 FT /product= "LGR6 polypeptide"
 FT
 XX WO200185768-A2.
 PN
 XX

PD 15-NOV-2001.
 XX
 PF 08-MAY-2001; 2001WO-US015002.
 XX
 PR 08-MAY-2000; 2000US-00566588.
 XX
 PA (MILL-) MILLENNIUM PHARM INC.
 XX
 PI Gu W;
 XX
 DR WPI: 2002-055584/07.
 DR P-PSDB; AAG66140.
 XX
 PT Novel isolated large G-protein coupled receptor 6 polypeptide, useful for
 PT diagnosing and treating weight disorder, metabolic disorder, central
 PT nervous system disorder, endocrine disorder and cardiovascular disorder.
 XX
 PS Claim 2; Fig 8; 198pp; English.
 XX
 CC The invention provides isolated large G-protein coupled receptor 6 (LGR6)
 CC polypeptides and polynucleotides encoding them. The LGR6 polypeptides and
 CC polynucleotides are useful as targets for developing modulating agents
 CC that regulate a variety of cellular process, e.g. neural and endocrine
 CC processes, as well as thermogenesis. They are useful for developing novel
 CC diagnostic and therapeutic agents for LGR6 associated disorders such as
 CC weight disorders (anorexia, obesity), eating disorders, acquired
 CC immunodeficiency syndrome (AIDS)-related wasting, cancer-related wasting,
 CC metabolic disorders, central nervous system disorders (e.g. Alzheimer's
 CC disease, Parkinson's disease, Huntington's disease, Gilles de la
 CC Tourette's syndrome, multiple sclerosis, amyotrophic lateral sclerosis,
 CC epilepsy, Jakob-Creutzfeldt disease, depression, schizophrenia,
 CC obsessive-compulsive disorder, anxiety, panic disorder, migraine),
 CC endocrine disorders (hypo and hyperthyroidism, dwarfism, gigantism,
 CC acromegaly), and cardiovascular disorders (atherosclerosis, restenosis,
 CC hypertension, myocardial infarction, arrhythmia). The LGR6 polypeptides,
 CC polynucleotides and antibodies are useful in screening assays, detection
 CC assays (e.g., chromosomal mapping, tissue typing, forensic biology),
 CC predictive medicine (e.g. diagnostic assays, prognostic assays,
 CC monitoring clinical trials and pharmacogenomics), and in methods of
 CC treatment (e.g. therapeutic and prophylactic). The present sequence
 CC represents a human LGR6 polypeptide coding sequence
 XX
 SQ Sequence 2208 BP; 396 A; 710 C; 631 G; 471 T; 0 U; 0 Other;
 Alignment Scores:
 Pred. No.: 1,31e-237 Length: 2208
 Score: 3045.50 Matches: 602
 Percent Similarity: 92.61% Conservative: 12
 Best Local Similarity: 90.80% Mismatches: 19
 Query Match: 92.09% Indels: 32
 DB: 6 Gaps: 4
 US-10-664-667-5 (1-633) x AAI67925 (1-2208)
 QY 1 AnThrThrHisTyr-----
 Db 224 AACTCCACACACTATCTCTGAATGTGCGCATCCAGAGTTCCAGATCTCAAAG 283
 QY 6 ---ArgGluSerTrpTyrAla-----CysArgTyrArg 15
 Db 284 GCACCCAGCGCTGAGATCTCTGACCCCTGACCCGCGCAGGATCCGGTGTCTCCATCGG 343
 QY 16 -----SerGlyIleProGlySer-ThrHisAlaSerValGluArgSerGln---G1 31
 Db 344 GGATGTGCCAAGCTGCCAGGCTCCGAGTCTCGAACTGCTCACAATCAATTGAGG 403
 QY 31 yLeuSerLeuProAlaHisProAlaSerLeuAlaAlaAlaSerAsnThrAl 51
 Db 404 AGCTGCCCGCGCTGCACAGGTGTCAAGAAATTGGAGAAA-TCGGCTCCCAACAACCGC 462
 QY 51 aserGlyLysLeuGlu***AspThrPheSerGlnLeuSerSerLeuGlnAlaLeuAspLe 71
 Db 463 ATCTGGAAATTGGAGCT-GACACCTTCAGCCAGCTGAGCTCCCTGCAAGCCCTGATCT 521

QY 71 uSerTrpAenAlaIleArgSerIleHisProGluAlaPheSerThrLeuHisSerLeuVa 91
 DB 522 TAGCTGGAAAGCCATCCGGTCCATCCACCTGGAGGCTTCTCCACCTGCACTCCCTGGT 581
 QY 91 llySerLeuAspLeuThrAspAsnGlnLeuThrThrLeuProLeuAlaGlyLeuGlyLe 111
 DB 582 CAAGCTGGACCTGACAGACAACAGCTGACACACTGCCCCCTGGCTGGAGTTGGGGGCTT 641
 QY 111 uMetHisLeuLeuLeuLeuGlyAsnLeuAlaLeuSerGlnAlaPheSerLysAspSerPh 131
 DB 642 GATGCATCTGAAGCTCAAAAGGAACTTGCTCTCTCCAGGCTTCTCCAAAGGACAGTTT 701
 QY 131 eProLysLeuArgIleLeuGluValProTyrAlaTyrGlnCysCysProTyrGlyMetCy 151
 DB 702 CCCAAACTGAGGATCTGGAGGTGCTTATGCCCTACCAAGTCTGCTCTATGGGATGTG 761
 QY 151 sAlaSerPhePheLysAlaSerGlyGlnTrpGluAlaGluAspLeuHisLeuAspAspG1 171
 DB 762 TGCCAGCTTCTTCAAGGCTCTGGGCAGTGGGAGGCTGAAGACCTTCACTTGTATGATGA 821
 QY 171 uGluSerSerLysArgProLeuGlyLeuLeuAlaArgGlnAlaGluAsnHisTyrAspG1 191
 DB 822 GGAAGTCTTCAAAAGGCTCTGGGCTCTCTTGCCAGACAAGCAGAGAACCACTATGACCA 881
 QY 191 nAspLeuAspGluLeuGlnLeuGluMetGluAspSerLysProHisProSerValGlnCy 211
 DB 882 GGACCTGGATGATCTCCAGCTGGAGTGGAGGACTCAAAAGCCACACCCAGTGTCCAGTG 941
 QY 211 sSerProThrProGlyProPheLysProCysGluTyrLeuPheGluSerTrpGlyLeAr 231
 DB 942 TAGCCCTACTCCAGGCTCTTCAAGCCTGTGAGTACCTCTTTGAAAGCTGGGGCATCG 1001
 QY 231 gLeuAlaValTrpAlaIleValLeuLeuSerValLeuCysAsnGlyLeuValLeuLeuTh 251
 DB 1002 CCTGGCCGTGTGGGCATCGTGTCTCTCCGTCTCTGCAATGAGCTGGTGTCTGAC 1061
 QY 251 rValPheAlaGlyGlyProAlaProLeuProValLysPheValValGlyAlaIleAl 271
 DB 1062 CGTGTTCGTGGGGGCTGCCCCCTGCCCCCGGTCAAGTTTGGTAGGTGGCATTCG 1121
 QY 271 aGlyAlaAsnThrLeuThrGlyIleSerCysGlyLeuLeuAlaSerValAspAlaLeuTh 291
 DB 1122 AGGCGCCCAACACTTGACCTGGCATTTCTGTGGCTTCTAGCCTCAGTCGATGCCCTGAC 1181
 QY 291 rPheGlyGlnPheSerGluTyrGlyAlaArgTrpGluThrGlyLeuGlyCysArgAlaTh 311
 DB 1182 CTTTGGTCAGTTCTCTGAGTACGAGCCGCTGGGAGACGGGGCTAGGCTGCCGGGCCAC 1241
 QY 311 rGlyPheLeuAlaValLeuGlySerGluAlaSerValLeuLeuLeuThrLeuAlaAlaVa 331
 DB 1242 TGCGTTCCTGGCAGTACTTGGGTGGAGGCAATCGGTGCTGCTGCTCACTCTGGCGCGAGT 1301
 QY 331 lGlnCysSerValSerValSerCysValArgAlaTyrGlyLysSerProSerLeuGlyLys 351
 DB 1302 GCAGTGCAGGCTCTCCGCTCTCTGTCTCGGGCTATGGGAAGTCCCCCTCCCTGGGAG 1361
 QY 351 rValArgAlaGlyValLeuGlyCysLeuAlaLeuAlaGlyLeuAlaAlaLeuProLe 371
 DB 1362 GCTTCGAGCAGGGGCTTCTAGGTGCTGGCACTGGCAGGGCTGGCGCGGCACTGCCCT 1421
 QY 371 uAlaSerValGlyGluTyrGlyAlaSerProLeuCysLeuProTyrAlaProProGluG1 391
 DB 1422 GGCCTCAGTGGAGAAATACGGGGCTTCCCTCTGCTGCTGCTGCTGCTGCTGCTGAGG 1481
 QY 391 yGlnProAlaAlaLeuGlyPheThrValAlaLeuValMetMetAsnSerPheCysPheLe 411
 DB 1482 TCAGCCAGCAGCCCTGGGCTTCCACCGTGGCCCTGGTGTATGATGAACCTCTTCTGTTCCT 1541
 QY 411 uValValAlaGlyAlaTyrIleLysLeuTyrCysAspLeuProArgGlyAspPheGluAl 431
 DB 1542 GGTCTGGCCGCTGCTCATCAACTGTACTGTGACCTGCGGGGCGGAGCTTTGAGGC 1601

QY 431 aValTrpAspCysAlaMetValArgHisValAlaTrpLeuIlePheAlaAspGlyLeuLe 451
 DB 1602 CGTGTGGAGCTGCGCCATGGTGGAGCAGCTGGCTGGCTCATCTTCGACAGCGGGCTCT 1661
 QY 451 uTyrCysProValAlaPheLeuSerPheAlaSerMetLeuGlyLeuPheProValThrPr 471
 DB 1662 CTACTGTCCGTGGCTTCTCTCAGCTTCCGCTCCTCAGCTTCCCTGTGTCAGGCC 1721
 QY 471 oGluAlaValLysSerValLeuLeuValValLeuProLeuProAlaCysLeuAsnProLe 491
 DB 1722 CGAGGCGCTCAAGTCTGCTGCTGGTGGTGGTGGCTGCCCTGCCCTGCCCTCAACCCACT 1781
 QY 491 uLeuTyrLeuLeuPheAsnProHisPheArgAspAspLeuArgArgLeuArgProArgAl 511
 DB 1782 GCTGTACTGCTCTTCAACCCCACTTCCGGGATGACTTCCGGGGCTTCGGCCCCGGCG 1841
 QY 511 aGlyAspSerGlyProLeuAlaTyrAlaAlaAlaGlyGluLeuGluLysSerSerCysAs 531
 DB 1842 AGGGGACTCAGGGGCCCTAGCCTATGCTGGCGCCGGGAGCTGGAGAAGAGCTCCTGTGA 1901
 QY 531 pSerThrGlnAlaLeuValAlaPheSerAspValAspLeuIleLeuGluAlaSerGluAl 551
 DB 1902 TTCTACCCAGGCCCTGGTAGCCTTCTCTGATGTGGATCTCATCTTGGAGCTTCTGAAGC 1961
 QY 551 aGlyArgProProGlyLeuGluThrTyrGlyPheProSerValThrLeuLysSerCysG1 571
 DB 1962 TGGCGGCCCTTGGGCTGGAGACTATGGCTTCCCTCAGTGACCTCATCTCTCTGTCA 2021
 QY 571 nGlnProGlyAlaProArgLeuGluGlySerHisCysValGluProGluGlyAsnHisPh 591
 DB 2022 GCAGCCAGGGGCCCCAGGCTGGAGGCGAGCCATTGTGTAGAGCAGAGGGGAACCACTT 2081
 QY 591 eGlyAsnProGlnProSerMetAspGlyGluLeuLeuLeuArgAlaGluGlySerThrPr 611
 DB 2082 TGGGAACCCCAACCCCTCCATGGATGGAGAACTGCTCTGAGGCGAGAGGATCTACGCC 2141
 QY 611 oAlaGlyGlyGlyLeuSerGlyGlyGlyPheGlnProSerGlyLeuAlaPheAlaSe 631
 DB 2142 AGCAGGTGGAGCTTGTTCAGGGGGTGGCGGCTTTCAGCCCTCTGGCTTGGCTTGGCTTC 2201
 QY 631 rHisVal 633
 DB 2202 ACACGTG 2208
 RESULT 10
 ID ADN02247 standard; cDNA; 2208 BP.
 AC ADN02247;
 XX 17-JUN-2004 (first entry)
 DT Human partial large G protein-coupled receptor, LGR6, CDS #2.
 XX Human; ss; large G protein-coupled receptor 6; LGR6; heart failure;
 XX ischaemic heart disease; myocardial infarction; hypertension;
 KW pericarditis; atherosclerosis; hepatic failure; viral hepatitis; cancer;
 KW intracerebral haemorrhage; brain abscess; Alzheimer's disease;
 KW pick disease; Parkinson's disease; Huntington's disease;
 KW amyotrophic lateral sclerosis; inflammatory bowel disease;
 KW Crohn's disease; benign prostatic hypertrophy;
 KW systemic lupus erythematosus.
 XX Homo sapiens.
 XX Key Location/Qualifiers
 CDS 1..2208
 FT /*tag= a
 FT /partial
 FT /note= "No start or stop codon shown"
 XX /product= "Partial LGR6"
 PN US2003166047-A1.

```
XX 04-SEP-2003.
XX
XX
XX 08-MAY-2001; 2001US-00851595.
XX
XX 06-MAY-1999; 99US-0132896P.
XX 08-MAY-2000; 2000US-00566588.
XX
XX (MILL-) MILLENNIUM PHARM INC.
XX
XX Gu W;
XX
XX WPI; 2003-898067/82.
XX P-PSDB; ADN02246.
XX
XX New isolated LGR6 nucleic acids and polypeptides, useful for diagnosing
XX and treating LGR6-mediated disorders, such as myocardial infarction,
XX hypertension, atherosclerosis, viral hepatitis, cancer and/or Alzheimer's
XX disease.
XX
XX Claim 1; SEQ ID NO 9; 145pp; English.
XX
XX PS The invention relates to an isolated LGR6 (large G protein-coupled
XX receptor) nucleic acid molecule comprising the mouse and human cDNA
XX sequences appearing as ADN02239, ADN02241, ADN02242, ADN02244, ADN02245,
XX ADN02247, ADN02248 and ADN02250. Also included are a host cell or a non-
XX human mammalian host cell containing an LGR6 nucleic acid (and producing
XX LGR6), an isolated LGR6 polypeptide appearing as ADN02240, ADN02243,
XX ADN02246 and ADN02249, an antibody which selectively binds to LGR6, a
XX method for detecting the presence of LGR6 in a sample, a kit (comprising
XX a compound which selectively binds to LGR6, or hybridises to LGR6 nucleic
XX acid, and instructions for use), a method for detecting LGR6 nucleic acid
XX in a sample, a method for identifying a compound which binds to LGR6, a
XX method for modulating the activity of LGR6 and a method for identifying a
XX compound that modulates the activity of LGR6. The methods and
XX compositions of the present invention are useful for diagnosing and
XX treating large G-protein coupled receptor 6 (LGR6)-mediated or -related
XX disorders, such as heart failure, ischaemic heart disease, myocardial
XX infarction, hypertension, pericarditis, atherosclerosis, hepatic failure,
XX viral hepatitis, cancer, intracerebral haemorrhage, brain abscess,
XX Alzheimer's disease, Pick disease, Parkinson's disease, Huntington's
XX disease, amyotrophic lateral sclerosis, inflammatory bowel disease,
XX Crohn's disease, benign prostatic hypertrophy and systemic lupus
XX erythematosus. The present sequence is a partial human cDNA for LGR6.
XX
XX SQ Sequence 2208 BP; 396 A; 710 C; 631 G; 471 T; 0 U; 0 Other;
XX
XX Alignment Scores:
XX Pred. No.: 1.31e-237 Length: 2208
XX Score: 3045.50 Matches: 602
XX Percent Similarity: 92.61% Conservative: 12
XX Best Local Similarity: 90.80% Mismatches: 19
XX Query Match: 92.09% Indels: 32
XX DB: 11 Gaps: 4
XX
XX US-10-664-667-5 (1-633) x ADN02247 (1-2208)
XX
XX QY 1 AenThrThrHisTyr----- 5
XX
XX DB 224 AACTCCACACTATCTCTGTAATGTGGCATGGACATCGAGAGTTTCCAGATCTCAAG 283
XX
XX QY 6 ---ArgGluSerTrpTyrAla-----CysArgTyrArg 15
XX
XX DB 284 GCACACCAGCCTGGAGATCTCTGACCTGACCGCGGAGGCATCGGCTGCCATCGG 343
XX
XX QY 16 -----SerGlyLeuProGlySer-ThrHisAlaSerValGluArgSerGln---G1 31
XX
XX DB 344 GGATGTGCCAACACAGCTGCCAGGCTCCGAGTCTCTGCAATCAATTTGAGG 403
XX
XX QY 31 yLeuSerLeuProAlaHisProAlaSerLeuAlaAlaLeuAlaAlaSerAsnThrAl 51
XX
XX DB 404 AGCTGCCAGCCTGCACAGGTGTCAAGAAATTGGAGAAA-TCGGGCTCCACACACCGC 462
```

```
QY 51 aSerGlyLysLeuGlu***AepThrPheSerGlnLeuSerSerLeuGlnAlaLeuAspLe 71
DB 463 ATCTGGGAAATTGGAGCT-GACACCTTCAGCCAGCTGAGCTCCCTGCAAGCCCTGATCT 521
QY 71 uSerTrpAsnAlaIleArgSerIleHisProGluAlaPheSerThrLeuHisSerLeuVa 91
DB 522 TAGCTGGAAACCCATCCGGTCCATCCACCTTGAGCCCTTCTCCACCCCTGCACTCCCTGGT 581
QY 91 lLysLeuAspLeuThrAspAsnGlnLeuThrThrLeuProLeuAlaGlyLeuGlyLe 111
DB 582 CAAGCTGGACCTGCACAGACCAACAGCTGACCACTGCCCTGGCTGGCTGGGGGCTT 641
QY 111 uMethHisLeuLysLeuLysGlyAsnLeuAlaLeuSerGlnAlaPheSerLysAspSerPh 131
DB 642 GATGCATCTGAAGCTCAAAAGGGAACCTTGTCTCTCCAGGCGCTTCTCAAGGACAGTTT 701
QY 131 eProLysLeuArgIleLeuGluValProTyrAlaTyrGlnCysCysProTyrGlyMetCy 151
DB 702 CCCAAAACCTGAGGATCTCGAGGTCCTTATGCCCTACCAAGTCTGCTCTTATGGGATGTG 761
QY 151 sAlaSerPhePheLysAlaSerGlyGlnTrpGluAlaGluAspLeuHisLeuAspAspG1 171
DB 762 TGCCAGCTTCTTCAAGGCCCTCTGGGCACTGGGAGCTGAAGACCTTCACTTGTATGATGA 821
QY 171 uGluSerSerLysArgProLeuGlyLeuLeuAlaArgGlnAlaGluAsnHisTyrAspG1 191
DB 822 GGAGTCTTCAAAAAGGCGCCCTGGGCTCTCTTGGCCAGACAAGAGAACCACTATGACCA 881
QY 191 nAspLeuAspGluLeuGlnLeuGluMetGluAspSerLysProHisProSerValGlnCy 211
DB 882 GGACTGTGATGAGCTCCAGCTGGAGATGGAGGACTCAAAGCCACACCCAGTGTCCAGTG 941
QY 211 sSerProThrProGlyProPheLysProCysGluTyrLeuPheGluSerTrpGlyIleAr 231
DB 942 TAGCCCTACTCCAGGCCCTTCAAGCCCTGTGAGTACCTCTTTGAAAGCTGGGGCATCG 1001
QY 231 gLeuAlaValTrpAlaIleValLeuLeuSerValLeuCysAsnGlyLeuValLeuLeuTh 251
DB 1002 CCGTGGCGTGTGGGCGCATCTGTTGTCTCTCGTGTCTGCAATGGAGCTGGTCTGCTGAC 1061
QY 251 rValPheAlaGlyGlyProAlaProLeuProProValLysPheValValGlyAlaIleAl 271
DB 1062 CGTGTCTCGTGGGCGGCTGCCCCCTGCCCCCGGTCAAGTTTGTGGTAGGTGGCATGTC 1121
QY 271 aGlyAlaAsnThrLeuThrGlyIleSerCysGlyLeuLeuAlaSerValAspAlaLeuTh 291
DB 1122 AGCGCCCAACACTTGTACTGGCATTTCTGTGGGCTTCTAGCCCTCAGTCTGATCGCTGAC 1181
QY 291 rPheGlyGlnPheSerGluTyrGlyAlaArgTrpGluThrGlyLeuGlyCysArgAlaTh 311
DB 1182 CTTTGGTCAAGTCTCTGAGTACGAGCCCGCTGGGAGAGCGGGCTAGGCTGCCGGGCCAC 1241
QY 311 rGlyPheLeuAlaValLeuGlySerGluAlaSerValLeuLeuLeuThrLeuAlaAlaVa 331
DB 1242 TGGCTTCTCGCAGTACTTGGGTCTGGAGGCATCGGTGCTGTGCTCACTCTGGCGCGAGT 1301
QY 331 lGlnCysSerValSerValSerCysValArgAlaTyrGlyLysSerProSerLeuGlySe 351
DB 1302 CGAGTGCAGCGCTCTCCGCTCTCTGTGTCGGGCGCTTATGGGAAGTCCCCCTCCCTGGG 1361
QY 351 rValArgAlaGlyValLeuGlyCysLeuAlaLeuAlaGlyLeuAlaAlaLeuProLe 371
DB 1362 CGTTCGAGAGGGGTCTTAGGCTGCTGGCACTGGCAGGGCTGGCGCGCGCACCTGCCCT 1421
QY 371 uAlaSerValGlyGluTyrGlyAlaSerProLeuCysLeuProTyrAlaProGluG1 391
DB 1422 GGCCTCAGTGGGAATAACGGGGCTCCCCACTCTGCCCTGCGCTACCGCGCCACCTGAGGG 1481
QY 391 yGlnProAlaAlaLeuGlyPheThrValAlaLeuValMetMetAsnSerPheCysPheLe 411
DB 1482 TCAGCCAGACCCCTGGGCTTCAAGTGGGCTTCAAGTGGGCTTCAAGTGGGCTTCTTCTTCT 1541
QY 411 uValValAlaGlyAlaTyrIleLysLeuTyrCysAspLeuProArgGlyAspPheGluAl 431
```



```
QY 31 yLeuSerLeuProAlaHisProAlaSerLeuAlaLeuAlaLaSerAsnThrAl 51
Db   :::          :::          :
404 AGCTGCCAGCTGCACAGGTGCAGAAATTCGAGAAA-TGGGCTCCAAACACACCGC 462

QY 51 aSerGlyValLeuGlu***AspThrPheSerGlnLeuSerLeuGlnAlaLeuAspLe 71
Db   |||          |||          |||
463 ATCTGGGAATTTGGAGCT-GACACCTTCAGCCAGCTGAGCTCCCTGCAAGCCCTGGATCT 521

QY 71 uSerTrpAsnAlaIleArgSerIleHisProGluAlaPheSerThrLeuHisSerLeuVa 91
Db   |||          |||          |||
522 TAGCTGGAAAGCCATCCGGTCCATCCACCTGAGGCTTCTCCACCTGCACCTCCCTGGT 581

QY 91 llyLeuAspLeuThrAspAsnGlnLeuThrThrLeuProLeuAlaGlyLeuGlyGlyLe 111
Db   |||          |||          |||
582 CAAGCTGGACCTGCACAGACCAACAGCTGACCACTGCCCCCTGGCTGGAGCTTGGGGCTT 641

QY 111 uMethIleuLeuValLeuGlyValLeuAlaLeuSerGlnAlaPheSerLysAspSerPh 131
Db   |||          |||          |||
642 GATGCATCTGAAGCTCAAGGGAACCTTGCTCTCTCCAGGCTTCTCCAGGACAGTTT 701

QY 131 eProLysLeuArgIleLeuGluValProTyrAlaTyrGlnCysCysProTyrGlyMetCy 151
Db   |||          |||          |||
702 CCCAAACTGAGGATCCTGGAGTGCTTATGCTTACCAGTGCTGCTCCCTATGGGATGTG 761

QY 151 sAlaSerPhePheLysAlaSerGlyGlnTrpGluAlaGluAspLeuHisLeuAspAspGl 171
Db   |||          |||          |||
762 TGCCAGCTTCTTCAAGGCTCTGGGAGTGGGAGCTGAAGACCTTCACTTGTATGATGA 821

QY 171 uGluSerSerLysArgProLeuGlyLeuLeuAlaArgGlnAlaGluAenHisTyrAspGl 191
Db   |||          |||          |||
822 GGAGTCTTCAAAAGGCCCTGGGCTCTTCCAGCAAGCAGAGAACCACTATGACCA 881

QY 191 nAspLeuAspGluLeuGlnLeuGluMetGluAspSerLysProHisProSerValGlnCy 211
Db   |||          |||          |||
882 GGACCTGATGATGCTCCAGCTGGAGATGGAGGACTCAAAAGCCACACCCAGTGTCCAGTG 941

QY 211 sSerProThrProGlyProPheLysProCysGluTyrLeuPheGluSerTrpGlyLeAr 231
Db   |||          |||          |||
942 TAGCCCTACTCCAGGCCCTTCAAGCCCTGTGAGTACCTCTTTGAAAGCTGGGGCATCCG 1001

QY 231 gLeuAlaValTrpAlaIleValLeuLeuSerValLeuCysAsnGlyLeuValLeuLeuTh 251
Db   |||          |||          |||
1002 CTTGGCCCTGTGGGCCATCTGTCTCTCCGTGCTCTGCAATGAGCTGGTCTCTCTGAC 1061

QY 251 rValPheAlaGlyGlyProAlaProLeuProProValLysPheValValGlyValaIleAl 271
Db   |||          |||          |||
1062 CGTGTTCGTGGCGGGCTGCCCCCTGCCCCCGGTCAAGTTTGTGGTAGGTGCATTTGC 1121

QY 271 aGlyAlaAsnThrLeuThrGlyIleSerCysGlyLeuLeuAlaSerValAspAlaLeuTh 291
Db   |||          |||          |||
1122 AGGCGCCCAACACTTGTGACTGGCATTTCTGTGGCTTCTAGCCTCAGTCGATGCCCTGAC 1181

QY 291 rPheGlyGlnPheSerGluTyrGlyAlaArgTrpGluThrGlyLeuGlyCysArgAlaTh 311
Db   |||          |||          |||
1182 CTTTGTGTCAGTTCTCTGAGTAGAGCCGCTGGGAGAGCGGGCTAGGCTGCGGGGCCAC 1241

QY 311 rGlyPheLeuAlaValLeuGlySerGluAlaSerValLeuLeuThrLeuAlaAlaVa 331
Db   |||          |||          |||
1242 TGCTTCTCTGGCAGTACTTGGGTTCGGGAGCATCGGTGCTGTGCTCACTCTGGCGCGAGT 1301

QY 331 lGlnCysSerValSerValSerCysValArgAlaTyrGlyLysSerProSerLeuGlySe 351
Db   |||          |||          |||
1302 GCAGTGCAGCGTCTCGGTCTCTGTGTCGGGCCCTATGGGAATCCCCCTCCCTCGGGCAG 1361

QY 351 rValArgAlaGlyValLeuGlyCysLeuAlaLeuAlaGlyLeuAlaAlaLeuProLe 371
Db   |||          |||          |||
1362 CGTTCGAGCAGGGGCTTCTAGGCTGCTGGCACTGGCAGGGCTGGCGCGCACTCCGCT 1421

QY 371 uAlaSerValGlyGlyTyrGlyValaSerProLeuCysLeuProTyrAlaProProGluGl 391
Db   |||          |||          |||
1422 GGCCTTCAGTGGGAGAAATACGGGGCTCCCCACTCTGCTGCTTACGCGCCACCTGAGGG 1481

QY 391 yGlnProAlaAlaLeuGlyPheThrValAlaLeuValMetMetAsnSerPheCysPheLe 411
```

```
Db 1482 TCAGCCAGCAGCCCTGGGCTTTCACCGTGGCCCTGGTGTATGATGAACCTCTTCTGTTCCT 1541
QY 411 uValValAlaGlyAlaTyrIleLysLeuTyrCysAspLeuProArgGlyAspPheGluAl 431
Db 1542 GGTGTGGCCCGGTGGCTACATCAAACTGTACTGTGACCTGCGGGGGCGACTTGGAGC 1601
QY 431 aValTrpAspCysAlaMetValArgHisValAlaTrpLeuIlePheAlaAspGlyLeuLe 451
Db 1602 CGTGTGGAGCTGCGCATGTGTGAGCAGCTGGCTGGCTCATCTTCGCAGACGGGCTCCT 1661
QY 451 uTyrCysProValAlaPheLeuSerPheAlaSerMetLeuGlyLeuPheProValThrPr 471
Db 1662 CTACTGTCCCTGGGCTTCTCAGCTTCCGCTCCTCACTGCTGGGCTCTTCCCTGTCAACGCC 1721
QY 471 oGluAlaValLysSerValLeuLeuValValLeuProLeuProAlaCysLeuAsnProLe 491
Db 1722 CGAGCCGCTCAAGTCTGTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1781
QY 491 uLeuTyrLeuLeuPheAsnProHisPheArgAspAspLeuArgArgLeuArgProArgAl 511
Db 1782 GCTGTACCTGCTTCAACCCCACTTCCGGGATGACCTTCGGGGCTTCGGCCCCGGGC 1841
QY 511 aGlyAspSerGlyProLeuAlaTyrAlaAlaGlyGluLeuGluLysSerSerCysAs 531
Db 1842 AGGGGACTCAGGGCCCCCTAGCTTATGCTGCGGCGGGAGCTGGAGAAAGAGCTCCTGTGA 1901
QY 531 pSerThrGlnAlaLeuValAlaPheSerAspValAspLeuIleLeuGluAlaSerGluAl 551
Db 1902 TTCTACCCAGGCCCTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1961
QY 551 aGlyArgProGlyLeuGluThrTyrGlyPheProSerValThrLeuIleSerCysGl 571
Db 1962 TGGGCGGCCCTTGGGCTGGAGACTATGCTTCCCTCAGTGACCTCATCTCTCTGTCA 2021
QY 571 nGlnProGlyAlaProArgLeuGluGlySerHisCysValGluProGluGlyAsnHisPh 591
Db 2022 GCAGCCAGGGGCCCCAGGCTGGAGGGCAGCATTTGTGTAGAGCCAGAGGGAACCACTT 2081
QY 591 eGlyAsnProGlnProSerMetAspGlyGluLeuLeuLeuArgAlaGluGlySerThrPr 611
Db 2082 TGGGAACCCCAACCCCTCATGGATGGAGAACTGCTGCTGAGGGCAGAGGATCTACGCC 2141
QY 611 oAlaGlyGlyLeuSerGlyGlyGlyPheGlnProSerGlyLeuAlaPheAlaSe 631
Db 2142 AGCAGTGGAGGCTTGTTCAGGGGTGGGGCTTTTCAGGCCCTTGGCTTGGCTTGTCTTC 2201
QY 631 rHisVal 633
Db 2202 ACACGTG 2208

RESULT 12
AAI67924
ID AAI67924 standard; cDNA; 2711 BP.
XX
AC AAI67924;
DT 13-MAR-2002 (first entry)
XX
DE Partial cDNA of human LGR6 polypeptide (clone fahr).
XX
KW Large G-protein coupled receptor 6; LGR6; GPCR; immunomodulator;
KW anti-HIV; cytosolic; norepinephrine; neuroprotective; antiparkinsonian;
KW anticonvulsant; neuroleptic; antidepressant; tranquilizer; antimigraine;
KW osteopathic; antiarteriosclerotic; vasodilator; hypotensive; cardiac;
KW antiarrhythmic; anorectic; gene therapy; human; ss.
OS Homo sapiens.
XX
FH Key
FT CDS
FT
FT /product= "LGR6 polypeptide"
FT
```

XX WO200185768-A2.
 XX 15-NOV-2001.
 XX 08-MAY-2001; 2001WO-US015002.
 XX 08-MAY-2000; 2000US-00566588.
 XX (MILL-) MILLENNIUM PHARM INC.
 XX Gu W;
 XX WPI; 2002-055584/07.
 XX P-P8DB; AAG66140.
 XX Novel isolated large G-protein coupled receptor 6 polypeptide, useful for
 PT diagnosing and treating weight disorder, metabolic disorder, central
 PT nervous system disorder, endocrine disorder and cardiovascular disorder.
 XX
 XX Claim 2; Fig 8; 198pp; English.
 XX The invention provides isolated large G-protein coupled receptor 6 (LGR6)
 CC polypeptides and polynucleotides encoding them. The LGR6 polypeptides and
 CC polynucleotides are useful as targets for developing modulating agents
 CC that regulate a variety of cellular process, e.g. neural and endocrine
 CC processes, as well as thermogenesis. They are useful for developing novel
 CC diagnostic and therapeutic agents for LGR6 associated disorders such as
 CC weight disorders (anorexia, obesity), eating disorders, acquired
 CC immunodeficiency syndrome (AIDS)-related wasting, cancer-related wasting,
 CC metabolic disorders, central nervous system disorders (e.g. Alzheimer's
 CC disease, Parkinson's disease, Huntington's disease, Gilles de la
 CC Tourette's syndrome, multiple sclerosis, amyotrophic lateral sclerosis,
 CC epilepsy, Jakob-Creutzfeldt disease, depression, schizophrenia,
 CC obsessive-compulsive disorder, anxiety, panic disorder, migraine),
 CC endocrine disorders (hypo and hyperthyroidism, dwarfism, gigantism,
 CC acromegaly), and cardiovascular disorders (atherosclerosis, restenosis,
 CC hypertension, myocardial infarction, arrhythmia). The LGR6 polypeptides,
 CC polynucleotides and antibodies are useful in screening assays, detection
 CC assays (e.g., chromosomal mapping, tissue typing, forensic biology),
 CC predictive medicine (e.g. diagnostic assays, prognostic assays,
 CC monitoring clinical trials and pharmacogenomics), and in methods of
 CC treatment (e.g. therapeutic and prophylactic). The present sequence
 CC represents a partial cDNA sequence of human LGR6 polypeptide
 XX
 XX Sequence 2711 BP; 530 A; 837 C; 740 G; 604 T; 0 U; 0 Other;
 SQ
 Alignment Scores:
 Pred. No.: 1.73e-237 Length: 2711
 Score: 3045.50 Matches: 602
 Percent Similarity: 92.61% Conservative: 12
 Best Local Similarity: 90.80% Mismatches: 19
 Query Match: 92.09% Indels: 32
 DB: 6 Gaps: 4
 US-10-664-667-5 (1-633) x AAI67924 (1-2711)
 QY 1 AsnThrThrHisTyr----- 5
 DB 224 AACTCCACACACTATCTCTGAATGGTGCCATCGACATCCAGGAGTTTCCAGATCTCAAG 283
 QY 6 ---ArgGluSerTyrTyrAla-----CysArgTyrArg 15
 DB 284 GCACACACAGCTGGAGATCTGACCTGACCCGCGCGGAGGATCGGCTGCTCCATCGG 343
 QY 16 -----SerGlyIleProGlySer-ThrHisAlaSerValGluArgSerGln---G1 31
 DB 344 GGATGTGCCAACAGCTGCCAGGCTCCGAGCTCTGGAACCTGCTCACAAATTTGAGG 403
 QY 31 YLeuSerLeuProAlaHisProAlaSerLeuAlaAlaAlaAlaSerAsnThrAl 51
 DB 404 AGCTGCCCGCCTGCACAGGTGTGAGAAATTTGGAGGAAA--TCGGCCTCCACACACCGC 462

QY 51 aSerGlyLeuLeuGlu***AspThrPheSerGlnLeuSerSerLeuGlnAlaLeuAspLe 71
 DB 463 ATCTGGGAAATTTGGAGCT-GACACCTTCAGCCAGCTGAGCTCCCTGCAAGCCCTGGATCT 521
 QY 71 uSerTrpAsnAlaIleArgSerIleHisProGluAlaPheSerThrLeuHisSerLeuVa 91
 DB 522 TAGCTGGAAGCCATCCGGTCCATCCACCTCGAGGCTTCTCCACCTCGACCTCCCTGGT 581
 QY 91 lLysLeuAspLeuThrAspAsnGlnLeuThrLeuProLeuAlaGlyLeuGlyGlyLe 111
 DB 582 CAAGCTGGACCTGACAGACCAACAGCTGACCACTGCCCTGGCTGGACTTTGGGGCTT 641
 QY 111 uMetHisLeuLysLeuLysGlyAsnLeuAlaLeuSerGlnAlaPheSerLysAspSerPh 131
 DB 642 GATGCACTCGAGCTCAAGGGAACTTGTCTCTCCAGGCTTCTCCAAGGACAGTTT 701
 QY 131 eProLysLeuArgIleLeuGluValProTyrAlaTyrGlnCysCysProTyrGlyMetCy 151
 DB 702 CCCAAAACCTGAGGATCTGGAGGTGCTTATGCTTACAGTGTCTGTCTATGGATGTG 761
 QY 151 sAlaSerPhePheLysAlaSerGlyGlnTyrGluAlaGluAspLeuHisLeuAspG1 171
 DB 762 TGCAGCTTCTTCAAGGCTCTTGGGAGTGGAGGCTGAAGACCTTCACTTGAATGA 821
 QY 171 uGluSerSerLysArgProLeuGlyLeuLeuAlaArgGlnAlaGluAsnHisTyrAspG1 191
 DB 822 GGAGTCTTCAAAAGGCTTGGGCTTCTTGCAGACAAAGCAGAGAACCACTATGACCA 881
 QY 191 nAspLeuAspGluLeuGlnMetGluAspSerLysProHisProHisProSerValGlnCy 211
 DB 882 GGACCTGGATGAGTCCAGCTGGAGATGGAGGATCAAGGCCACACCCCACTGTGCCAGTG 941
 QY 211 sSerProThrProGlyProPheLysProCysGluTyrLeuPheGluSerTyrGlyIleAr 231
 DB 942 TAGCCCTTACTCCAGGCTTCAAGGCTTGTGAGTACTCTTTGAAAGCTGGGGCATCCG 1001
 QY 231 gLeuAlaValTrpAlaIleValLeuLeuSerValLeuCysAsnGlyLeuValLeuLeuTh 251
 DB 1002 CTGGCCGCTGGGCCATCGTGTGCTCTCGTGCTCTGCAATGAGTGGTGTCTGTGAC 1061
 QY 251 rValPheAlaGlyCylProAlaProLeuProProValLysPheValValGlyAlaIleAl 271
 DB 1062 CGTGTTCGCTGGGGGCTGCTCCCTCCCGCCGCTCAAGTTTGTGTAGTGGATGTC 1121
 QY 271 aGlyAlaAsnThrLeuThrGlyIleSerCysGlyLeuLeuAlaSerValAlaLeuTh 291
 DB 1122 AGCGCCACACACCTTGACTGGCATTTCTGTGGCTTCTAGCTCAGTCGATGCCCTGAC 1181
 QY 291 rPheGlyGlnPheSerGluTyrGlyAlaArgTyrGluThrGlyLeuGlyCysArgAlaTh 311
 DB 1182 CTTTGTGTCAGTTCTCTGAGTACGAGGCCGCTGGAGACGGGGCTAGGCTGCCGGGCCAC 1241
 QY 311 rGlyPheLeuAlaValLeuGlySerGluAlaSerValLeuLeuLeuThrLeuAlaIaVa 331
 DB 1242 TGGCTTCTCGCAGTACTTGGGTCCGAGGATCGGTGCTGTCTGTCTGTCTGTGGCCGAGT 1301
 QY 331 lGlnCysSerValSerValSerCysValArgAlaTyrGlyLysSerProSerLeuGlySe 351
 DB 1302 GCAGTGGAGGCTTCCGCTCTCTGTGTCCGGGCTATGGGAAGTCCCTCCCTCGGGCAG 1361
 QY 351 rValArgAlaGlyValLeuGlyCysLeuAlaLeuAlaGlyLeuAlaAlaLeuProLe 371
 DB 1362 CGTTCGAGCAGGGGTCTTAGGCTGCTGGCAGTGGCAGGCTGGCCGCGCAGCTGCCCT 1421
 QY 371 uAlaSerValGlyGluTyrGlyAlaSerProLeuCysLeuProTyrAlaProProGluGl 391
 DB 1422 GGCTTCAGTGGGGAATACGGGGCTCCCTCCACTGTGCTGCCCTACCGCCACCTGAGGG 1481
 QY 391 YGlnProAlaAlaLeuGlyPheThrValAlaLeuValMetMetAsnSerPheCysPheLe 411
 DB 1482 TCAGCCAGAGCCTGGGCTTCCCGTGGCCCTGGTGTATGATGACTCTCTCTGTCTTCT 1541
 QY 411 uValValAlaGlyAlaTyrIleLysLeuTyrCysAspLeuProArgGlyAspPheGluAl 431

Db 1542 GGTCTGTGCGGTGCTTACATCAAACTGTACTGTGACCTGCGCGGGCGAGCTTTGAGGC 1601
Qy 431 aValTrpAapCysAlaMetValArGHisValAlaTrpLeuIlePheAlaAapGlyLeuLe 451
Db 1602 CGTGTGGAGACTGCGGCATGGTGGAGCAGCTGGCCCTGGCTCATCTTCGAGAGCGGCTCT 1661
Qy 451 uTyrCysProValAlaPheLeuSerPheAlaSerMetLeuGlyLeuPheProValThrPr 471
Db 1662 CTACTGTCCCGTGGGCTTCTCTCAGCTTCGCTCCATGCTGGGCGCTTCTCCCTGTCAAGCC 1721
Qy 471 oGluAlaValIysSerValLeuValValLeuProLeuProAlaCysLeuAenProLe 491
Db 1722 CGAGGCCGTCAAGTGTCTGCTGTGTGTGTGCTGCTGCTGCTGCTCAACCCACT 1781
Qy 491 uLeuTyrLeuLeuPheAenProHisPheArgAapAspLeuArgLeuAtpProArgAl 511
Db 1782 GCTGTACTGTCTTCTTAACCCCACTTCGGGATGACCTTCGGCGGCTTCGGCCCCCGGC 1841
Qy 511 aGlyAapSerGlyProLeuAlaTyrAlaAlaAlaGlyGluLeuGluLysSerSerCysAs 531
Db 1842 AGGGGACTCAGGGCCCTAGCTATGCTGCGGCGGGAGCTGGAGAGAGCTCTGTGA 1901
Qy 531 pSerThrGlnAlaLeuValAlaPheSerAapValAspLeuIleLeuGluAlaSerGluAl 551
Db 1902 TTCTACCCAGGCGCTGGTAGCTTCTCTGATGTGGATCTCATCTTGAAGCTTCTGAAGC 1961
Qy 551 aGlyArgProProGlyLeuGluThrTyrGlyPheProSerValThrLeuIleSerCysGl 571
Db 1962 TGGCGGGCCCCCTGGGCTGGAGACCTATGGCTTCCCTCAGTAGACCTCATCTCTGTCA 2021
Qy 571 nGlnProGlyAlaProArgLeuGluGlySerHisCysValGluProGluGlyAsnHisPh 591
Db 2022 GCAGCCAGGGGCCCCAGGCTGGAGGCGAGCCATTGTGTAGAGCCAGAGGGGAACCACTT 2081
Qy 591 eGlyAenProGlnProSerMetAepGlyCyluLeuLeuAlaGluGlySerThrPr 611
Db 2082 TGGGAACCCCAACCTCCATCGATGGAGAACTGTGCTGTGGGGCAGAGGATCTACGCC 2141
Qy 611 oAlaGlyGlyLeuSerGlyGlyGlyPheGlnProSerGlyLeuAlaPheAlase 631
Db 2142 AGCAGGTGAGGCTTGTGAGGGGTGGCGGCTTTTCAGCCCTCTGGCTTGGCTTGGCTTC 2201
Qy 631 rHisVal 633
Db 2202 ACACGTG 2208

RESULT 13
ADN02245
ID ADN02245 standard; cDNA; 2711 BP.
XX AC ADN02245;
XX DT 17-JUN-2004 (first entry)
XX DE Human partial large G protein-coupled receptor, LGR6, cDNA #2.
XX KW Human; ss; large G protein-coupled receptor 6; LGR6; heart failure;
KW ischaemic heart disease; myocardial infarction; hypertension;
KW pericarditis; atherosclerosis; hepatic failure; viral hepatitis; cancer;
KW intracerebral haemorrhage; brain abscess; Alzheimer's disease;
KW Pick disease; Parkinson's disease; Huntington's disease;
KW amyotrophic lateral sclerosis; inflammatory bowel disease;
KW Crohn's disease; benign prostatic hypertrophy;
KW systemic lupus erythematosus.
XX OS Homo sapiens.
XX FH Key
XX CDS Location/Qualifiers
FT 1..2211
FT /*tag= a
FT /partial
FT /note= "No start codon shown"

FT XX /product= "Partial LGR6"
PN US2003166047-A1.
XX 04-SEP-2003.
XX PF 08-MAY-2001; 2001US-00851595.
XX PR 06-MAY-1999; 99US-0132896P.
XX PR 08-MAY-2000; 2000US-00566588.
XX (MILL-) MILLENNIUM PHARM INC.
XX Gu W;
XX WPI; 2003-898067/82.
XX P-PSDB; ADN02246.
XX New isolated LGR6 nucleic acids and polypeptides, useful for diagnosing
PT and treating LGR6-mediated disorders, such as myocardial infarction,
PT hypertension, atherosclerosis, viral hepatitis, cancer and/or Alzheimer's
PT disease.
XX Claim 1; SEQ ID NO 7; 145pp; English.
XX The invention relates to an isolated LGR6 (large G protein-coupled
CC receptor) nucleic acid molecule comprising the mouse and human cDNA
CC sequences appearing as ADN02239, ADN02241, ADN02242, ADN02244, ADN02245,
CC ADN02247, ADN02248 and ADN02250. Also included are a host cell or a non-
CC human mammalian host cell containing an LGR6 nucleic acid (and producing
CC LGR6), an isolated LGR6 polypeptide appearing as ADN02240, ADN02243,
CC ADN02246 and ADN02249, an antibody which selectively binds to LGR6, a
CC method for detecting the presence of LGR6 in a sample, a kit (comprising
CC a compound which selectively binds to LGR6, or hybridises to LGR6 nucleic
CC acid, and instructions for use), a method for detecting LGR6 nucleic acid
CC in a sample, a method for identifying a compound which binds to LGR6, a
CC method for modulating the activity of LGR6 and a method for identifying a
CC compound that modulates the activity of LGR6. The methods and
CC compositions of the present invention are useful for diagnosing and
CC treating large G-protein coupled receptor 6 (LGR6)-mediated or -related
CC disorders, such as heart failure, ischaemic heart disease, myocardial
CC infarction, hypertension, pericarditis, intracerebral haemorrhage, brain abscess,
CC viral hepatitis, cancer, Parkinson's disease, Huntington's
CC Alzheimer's disease, Pick disease, Parkinson's disease, Huntington's
CC disease, amyotrophic lateral sclerosis, inflammatory bowel disease,
CC Crohn's disease, benign prostatic hypertrophy and systemic lupus
CC erythematosus. The present sequence is a partial human cDNA for LGR6.
XX SQ Sequence 2711 BP; 530 A; 837 C; 740 G; 604 T; 0 U; 0 Other;
Alignment Scores:
Pred. No.: 1-73e-237 Length: 2711
Score: 3045.50 Matches: 602
Percent Similarity: 92.61% Conservative: 12
Best Local Similarity: 90.80% Mismatches: 19
Query Match: 92.09% Indels: 32
Dbs: 11 Gaps: 4
US-10-664-667-5 (1-633) x ADN02245 (1-2711)
Qy 1 AsnThrThrHisTyr----- 5
Db 224 AACCCACACACTATCTCTGAATGTGCGATCCAGATCCAGAGTTTCCAGATCTCAAG 283
Qy 6 ---ArgGluSerTrpTyrAla-----CysArgTyrArg 15
Db 284 GCACCACAGCCTGGAGATCTGACCCCTGACCCCGCGAGGATCCGGCTGCTCCATCGG 343
Qy 16 -----SerGlyIleProGlySer-ThrHisAlaSerValGluArgSerGln---G1 31
Db 344 GGATGTGCCAACAGCTGCCCGAGGCTCCGAGTCTCGGAACTGTCTCACATCAATGAGG 403
Qy 31 yLeuSerLeuProAlaHisProAlaSerLeuAlaAlaLeuAlaAlaSerAsnThrThrAl 51

/product= "LGR6 polypeptide #2"
/note= "This sequence lacks a start codon"

US2004058377-A1.
25-MAR-2004.
18-SEP-2003; 2003US-00664667.
06-MAY-1999; 99US-0132896P.
08-MAY-2000; 2000US-0056588.
08-MAY-2001; 2001US-00851595.
(MILL-) MILLENNIUM PHARM INC.
Gu W;
WPI; 2004-268789/25.
P-PSDB; ADK19411.
New large G-protein coupled receptor 6 nucleic acid molecules and polypeptides, useful for diagnosing, preventing or treating diseases associated with aberrant nucleic acid or protein activity, e.g. obesity, anorexia or cachexia.
Claim 2; SEQ ID NO 7; 145pp; English.
The present invention relates to the isolation of novel members of the G-protein coupled receptor (GPCR) family designated as large G-protein coupled receptor 6 (LGR6), and the polynucleotide sequences encoding them. The invention also discloses LGR6 fusion proteins, antigenic peptides, anti-LGR6 antibodies, recombinant expression vectors, host cells, and non-human transgenic animals in which an LGR6 gene has been introduced or disrupted. The LGR6 polynucleotide sequences, polypeptide sequences, and antibodies are useful for diagnosing, treating or preventing diseases associated with aberrant LGR6 expression or activity, such as weight disorders (e.g. obesity, anorexia nervosa or cachexia), neural disorders (e.g. central nervous system disorder, including Alzheimer's disease, Parkinson's disease, AIDS-related dementia, multiple sclerosis and Huntington's disease), endocrine disorders (e.g. hypothyroidism, hyperthyroidism or acromegaly), or cardiovascular disorders (e.g. ischaemia-reperfusion injury, coronary artery disease, myocardial infarction, arrhythmia, atherosclerosis, hypertension or congestive heart failure). The LGR6 polynucleotide and polypeptide sequences are also useful as targets for developing modulating agents that regulate a variety of cellular processes, e.g. neural and endocrine processes and thermogenesis. They can be used in screening assays (e.g. chromosome mapping, tissue typing or in forensic biology), or in predictive medicine (e.g. diagnostic assays, prognostic assays, monitoring clinical trials or pharmacogenomics). The present sequence encodes a human LGR6 polypeptide.
XX Sequence 2711 BP; 530 A; 837 C; 740 G; 604 T; 0 U; 0 Other;
SQ
Alignment Scores:
Pred. No.: 1,73e-237 Length: 2711
Score: 3045.50 Matches: 602
Percent Similarity: 92.61% Conservatives: 12
Best Local Similarity: 90.80% Mismatches: 19
Query Match: 92.09% Indels: 32
DB: 12 Gaps: 4
US-10-664-667-5 (1-633) x ADK19410 (1-2711)
QY 1 AenThrThrHisTyr----- 5
DB 224 AACTCCACACATATCTCTGAATGGTGGCCATGGACATCAGGATTTCCAGATCTCAAAG 283
QY 6 ---ArgGluSerTrpTyrAla-----CysArgTyrArg 15
DB 284 GCACACACGCTGAGATCTTGACCTGACCCGCGAGGATCCGGCTGCTCCCATCGG 343
QY 16 -----SerGlyIleProGlySer-ThrHisAlaSerValGluArgSerGln---G1 31

Db 344 GGATGTGCCAACAGCTGCCAGGCTCCGAGTCTCGAACTGCTCTCACAATCAAAATTGAGG 403
QY 31 yLeuSerLeuProAlaHisProAlaSerLeuAlaAlaLeuAlaHisSerAenThrThAl 51
Db 404 AGCTGCCAGCCTGTCACAGGTGTGAGAAATTTGGAGGAAA-TCGGGCTCCAAACACACGCG 462
QY 51 aSerGlyLysLeuGlu***AspThrPheSerGlnLeuSerSerLeuGlnAlaLeuAspLe 71
Db 463 ATCTGGGAAATTTGGAGCT-GACACCTTCAGCCAGCTGAGCTCCCTGCAAGCCCTGGAICT 521
QY 71 uSerTrpAenAlaIleArgSerIleHisProGluAlaPheSerThrLeuHisSerLeuVa 91
Db 522 TAGCTGGAAACGCTCCGCTCCATCCACCTGAGGCTTCTCCACCTGCACTCCCTGCT 581
QY 91 llyLeuAspLeuThrAspAenGlnLeuThrLeuProLeuAlaGlyLeuGlyLe 111
Db 582 CAAGCTGGACCTGACAGACAAACAGCTGACCACTGCCCTGGCTGGCTGGGGGCTT 641
QY 111 uMetHisLeuLysLeuLysGlyAenLeuAlaLeuSerGlnAlaPheSerLysAspSerPh 131
Db 642 GATGATCTGAAGCTCAAAGGAACTTGCTCTCTCCAGGCTTCTCCAGGACAGATT 701
QY 131 eProLysLeuArgIleLeuGluValProTyrAlaTyrGlnCysCysProTyrGlyMetCy 151
Db 702 CCCAAACTGAGGATCTCGAGGTGCTTATGCTTACCTGCTGCTGCTGCTGCTGCTG 761
QY 151 aAlaSerPhePheLysAlaSerGlyGlnTrpGluAlaGluAspLeuHisLeuAspG 171
Db 762 TGCCAGCTTCTTCAAGGCTCTGGGAGCTGAGGAGCTGAAGACCTTCACTTGATGATGA 821
QY 171 uGluSerSerLysArgProLeuGlyLeuAlaArgGlnAlaGluAenHisTyrAspG 191
Db 822 GGAGTCTTCAAAAGGCCCTTGGGCTCTTGGCCAGACAGACAGACACCTATGACCA 881
QY 191 nAspLeuAspGluLeuGlnLeuGluMetGluAspSerLysProHisProSerValGlnCy 211
Db 882 GGACCTGGATGAGTCTCAGCTGGAGAGTGGAGAGCTCAAAGCCACACCCAGTGTCCAGTG 941
QY 211 sSerProThrProGlyProPheLysProCysGluTyrLeuPheGluSerTrpGlyLea 231
Db 942 TAGCCCTACTCCAGGCCCTTCAAGCCCTTGAGTACCTCTTTGAAAGCTGGGGCATCG 1001
QY 231 gLeuAlaValTrpAlaIleValLeuLeuSerValLeuCysAenGlyLeuValLeuLeuTh 251
Db 1002 CTGGCCCTGTGGGCCATCGTGTGCTCTCGTGTCTGCAATGGACTGGTGTGCTGCTGAC 1061
QY 251 rValPheAlaGlyGlyProAlaProLeuProProValLysPheValValGlyAlaIleAl 271
Db 1062 CGTGTGCTGGCGGCTGCCCCCTGCCCGGTCAGTTTGTGGTAGGTGCGATTGC 1121
QY 271 aGlyAlaAenThrLeuThrGlyIleSerCysGlyLeuLeuAlaSerValAspAlaLeuTh 291
Db 1122 AGCGCCCAACACCTTGACTGGCATTTCTGTGGGCTTCTAGCCCTCAGTCGATGCCCTGAC 1181
QY 291 rPheGlyGlnPheSerGluTyrGlyAlaArgTrpGluThrGlyLeuGlyCysArgAlaTh 311
Db 1182 CTTTGGTCAGTCTCTGATACGGAGCCCGCTGGAGAGCGGGGCTAGGCTGCCGGGCCAC 1241
QY 311 rGlyPheLeuAlaValLeuGlySerGluAlaSerValLeuLeuLeuThrLeuAlaAlaVa 331
Db 1242 TGGCTCTCTGGCAGTACTTGGGTGGAGGATCGGTGCTGCTGCTGCTGCTGCTGCTGCTG 1301
QY 331 lGlnCysSerValSerValSerCysValArgAlaTyrGlyLysSerProSerLeuGlySe 351
Db 1302 CAGTGGACGCTCTCCGCTCTCTGTGTCGGGCTTATGGGAAGTCCCCCTCCCTGGGAG 1361
QY 351 rValArgAlaGlyValLeuGlyCysLeuAlaLeuAlaGlyLeuAlaAlaLeuProLe 371
Db 1362 GTTTCGACAGGGGTCTTAGGCTGCTGAGGCTGAGGAGGCTGGCGGCGGCGGCTCCCT 1421
QY 371 uAlaSerValGlyGluTyrGlyAlaSerProLeuCysLeuProTyrAlaProProGluGl 391

Db 1422 GGCCTCAGTGGGAGAAATACGGGGCTCCCACTCTGCCTGCCTACGGCCACCTGAGGG 1481
 Qy 391 yGlnProAlaLeuGlyPheThrValAlaLeuValMetMetAsnSerPheCysPheLe 411
 Db 1482 TCAGCCAGCAGCCCTGGGCTTACCGTGGCCCTGGTGATGATGAACCTCTCTGTTCCT 1541
 Qy 411 uValValAlaGlyAlaTyrlleLysLeuTyrcysAspLeuProArgGlyAspPheGluAl 431
 Db 1542 GGTCTGCGCGGTGCTACATCAACTACTGTGACCTGCGCGGGCGACTTTGAGGC 1601
 Qy 431 aValTrpAspCysAlaMetValArgHisValAlaTrpLeuIlePheAlaAspGlyLeuLe 451
 Db 1602 CGTGTGGGACTGCGGCATGGTGAGGACGCTGGCCCTGCTCATCTTCGCAGACGGGCTCCT 1661
 Qy 451 uTyrcysProValAlaPheLeuSerPheAlaSerMetLeuGlyLeuPheProValThrPr 471
 Db 1662 CTACTGTCCCGTGGCTTCTCCTCAGCTTCGCTCCATGCTGGGCCCTCTTCCTGTACGCC 1721
 Qy 471 oGluAlaValLysSerValLeuLeuValValLeuProLeuProAlaCysLeuAsnProLe 491
 Db 1722 CGAGGCCCTCAAGTCTGCTGCTGGTGCTGCTGCCCTGCTGCTGCCCTCAACCACT 1781
 Qy 491 uLeuTyrlleLeuPheAsnProHisPheArgAspAspLeuArgLeuArgProArgAl 511
 Db 1782 GCTGTACTGTCTTCAACCCCACTTCGGGATGACCTTCGGGGCTTCGGCCCGCGC 1841
 Qy 511 aGlyAspSerGlyProLeuAlaTyrlleAlaAlaGlyLylLeuGlyLysSerCysAs 531
 Db 1842 AGGGGACTCAGGGCCCTTAGCTATGCTGCGCGCGGGAGCTGGAGAAAGCTCCTGTGA 1901
 Qy 531 pSerThrGlnAlaLeuValAlaPheSerAspValAspLeuIleLeuGluAlaSerGluAl 551
 Db 1902 TTCTACCCAGGCCCTGGTAGCTTCTGATGTGGATCTCAATCTGGAGGCTTCTGAAGC 1961
 Qy 551 aGlyArgProGlyLeuGluThrTyrcysPheProSerValThrLeuIleSerCysGl 571
 Db 1962 TGGCGGGCCCTGGGCTGGAGACCTATGCTTCCCTCAGTGACCTCACTCTCTGTGCA 2021
 Qy 571 nGlnProGlyAlaProArgLeuGlyLysSerHisCysValGluProGlyLysAsnHisPh 591
 Db 2022 GCAGCGAGGGGCCCGCGCTGGAGGGCAGCCATTTGTGTAGAGCCAGAGGGGAACCACTT 2081
 Qy 591 eGlyAsnProGlnProSerMetAspGlyGlyLeuLeuLeuArgAlaGlyLysSerThrPr 611
 Db 2082 TGGGAACCCCAACCTCCATGGATGGAGAACTGCTGTGAGGGCAGAGGATCTACGCC 2141
 Qy 611 oAlaGlyGlyGlyLeuSerGlyGlyGlyPheGlnProSerGlyLeuAlaPheAlaSe 631
 Db 2142 AGCAGGTGGAGGCTTGTTCAGGGGGTGGCGGCTTTTCAGGCCCTCTGGCTTGGCTTTC 2201
 Qy 631 rHisVal 633
 Db 2202 ACACGTG 2208

US-10-664-667-5 (1-633) x ADC16710 (1-2786)
 Qy 1 AsnThrThrHisTyr----- 5
 Db 802 AACTCCACACTATCTCGAATGGTGCATGACATCCAGGAGTTTCCAGATCTCAAG 861
 Qy 6 ---ArgGluSerTrpTyrAla----- 15
 Db 862 GCACCACGCTGGAGATCTTGACCTGACCCCGCAGGAGATCCGGCTGCTCCATCGG 921
 Qy 16 -----SerGlyIleProGlySer-ThrHisAlaSerValGluArgSerGln---G1 31
 Db 922 GGATGTCACACAGCTGCCAGGCTCGAGTCTCGGAACTGTCTCACAATCAATTTGAGG 981
 Qy 31 yLeuSerLeuProAlaHisProAlaSerLeuAlaLeuAlaLeuAlaSerAsnThrAl 51
 Db 982 AGCTGCCCGCTGCACAGGTGTGAGAAATTTGGAGGAAA-TCGGCCTCCACACACGC 1040
 Qy 51 aSerGlyLysLeuGlu***AspThrPheSerGlnLeuSerSerLeuGlnAlaLeuAspLe 71
 Db 1041 ATCTGGGAAATTGGAGCT-GACACCTTCAGCCAGCTGAGCTCCCTGCAAGCCCTGGATCT 1099
 Qy 71 uSerTrpAsnAlaIleArgSerIleHisProGluAlaPheSerThrLeuHisSerLeuVa 91

Alignment Scores:
 Pred. No.: 1,79e-237 Length: 2786
 Score: 3045.50 Matches: 602
 Percent Similarity: 92.61% Conservative: 12
 Best Local Similarity: 90.80% Mismatches: 19
 Query Match: 92.09% Indels: 32
 DB: 10 Gaps: 4

/partial
 /product= "GPCR protein"
 /note= "Start and stop codons are absent"

WO2003040371-A1.
 15-MAY-2003.
 05-NOV-2002; 2002WO-JP011485.
 06-NOV-2001; 2001JP-00340189.
 31-MAY-2002; 2002JP-00159448.
 (TAKE) TAKEDA CHEM IND LTD.
 Ikeda N, Miwa M, Ito T, Ohtaki T;
 WPI; 2003-441575/41.
 P-PSDB; ADC16711.
 G-protein coupled receptor protein for treatment of infection and cancer etc.

Disclosure; Page 139-141; 153pp; Japanese.

This invention relates to novel cDNA sequences encoding the human G-protein coupled receptor (GPCR) proteins known as TGR41, namely TGR41A, TGR41V, TGR41A2 and TGR41V2. Specifically, it refers to the recombinant DNA vectors, the antibodies against the novel proteins as well as their ligands, a screening method for the detection compounds that affect GPCR protein binding, and also the resultant diagnostic drugs. The present invention describes these compounds as antimetabolites, neuroprotective, cytosstatic, antiinflammatory, osteopathic and antibacterial. As such, through using gene therapy they can be useful in the treatment of disorders associated with the central nervous system, endocrine system, metabolism, inflammation, circulation, respiration, digestion, immune system, bone, cartilage, urinary system, transplantation, infection and cancer. This polynucleotide is the DNA (SeqID 20) encoding a human GPCR protein related to the exemplification of the invention.

Sequence 2786 BP; 507 A; 917 C; 777 G; 585 T; 0 U; 0 Other;

RESULT 15
 ID ADC16710 standard; DNA; 2786 BP.
 XX AC ADC16710;
 XX AC
 XX 18-DEC-2003 (first entry)
 DX Human G-protein coupled receptor DNA (SeqID 20).
 XX human; G-protein coupled receptor; GPCR; TGR41; antimetabolite;
 KW neuroprotective; cytosstatic; antiinflammatory; osteopathic;
 KW antibacterial; gene therapy; infection; cancer; gene; ds.
 XX

OS Homo sapiens.
 XX
 XX Key Location/Qualifiers
 FH 3..2786
 FT CDS
 FT /*tag= a

Db	1100	TAGCTGGAACGCCATCCGGTCCATCCACCCCTGAGGCTTCTCCACCCCTGCATCCCTGGT	1159
Qy	91	llyLeuAspLeuThrAspAsnGlnLeuThrThrLeuProLeuAlaGlyLeuGlyGlyLe	111
Db	1160	CAAGCTGGACCTGACAGACAACACAGCTGACACACATGCCCTGGCTGGACTTGGGGCTT	1219
Qy	111	uMethisLeuLysLeuLysGlyAsnLeuAlaLeuSerGlnAlaPheSerLysAspSerPh	131
Db	1220	GATCATCTGAAGCTCAAGAGGAACCTTGCTCTCTCCAGGCTTCTCCAGGACAGTTT	1279
Qy	131	eProLysLeuArgGlyLeuGluValProTyrAlaTyrGlnCysCysProTyrGlyMetCy	151
Db	1280	CCCAAACTGAGGATCCCTGGAGGTGCTTATCCCTACCAAGTCTGCTCCCTATGGATGTG	1339
Qy	151	sAlaSerPhePheLysAlaSerGlyGlnTrpGluAlaGluAspLeuHisLeuAspAspG1	171
Db	1340	TGCCAGCTTCTCAAGGCTCTGGGCGAGTGGAGGCTGAAGACCTTCACCTTGATGATGA	1399
Qy	171	uGluSerSerLysArgProLeuGlyLeuLeuAlaArgGlnAlaGluAsnHisTyrAspG1	191
Db	1400	GGAGTCTTCAAAAGGCCCTGGGCTCTTCCAGACAAGCAGAGAACCACTATGACCA	1459
Qy	191	nAspLeuAspGluLeuGlnLeuGluMetGluAspSerLysProHisProSerValGlnCy	211
Db	1460	GGACCTGGATGAGCTCCAGCTGGAGATGGAGGACTCAAGGCCACACCCCAAGTGCAGTG	1519
Qy	211	sSerProThrProGlyProPheLysProCysGluTyrLeuPheGluSerTrpGlyLear	231
Db	1520	TAGCCCTACTCCAGGCCCTTCAAGCCCTGTGAGTACTCTTTGAAAGCTGGGGCATCCG	1579
Qy	231	gLeuAlaValTrpAlaIleValLeuLeuSerValLeuCysAsnGlyLeuValLeuLeuTh	251
Db	1580	CTGGCCGTGTGGCCATCGTGTGCTCTCGTGTCTGCAATGGACTGGTCTGCTGAC	1639
Qy	251	rValPheAlaGlyGlyProAlaProLeuProProValLysPheValValGlyAlaIleAl	271
Db	1640	CGTGTTCGTGGCGGCTGCCCCCTGCCCCCGTCAAGTTTGTGTAGTGCATTTGC	1699
Qy	271	aGlyAlaLeuThrLeuThrGlyLysSerCysGlyLeuLeuAlaSerValAspAlaLeuTh	291
Db	1700	AGGCCCAACACCTTGACTGGCATTTCTGTGGCTTCTAGGCTCAGTGGATGCCCTGAC	1759
Qy	291	rPheGlyGlnPheSerGluTyrGlyAlaArgTrpGluThrGlyLeuGlyCysArgAlaTh	311
Db	1760	CTTTGGTCACTTCTGAGTACGGAGCCGCTGGAGAGCGGGCTAGGCTGCCGGGCCAC	1819
Qy	311	rGlyPheLeuAlaValLeuGlySerGluAlaSerValLeuLeuLeuThrLeuAlaAlaVa	331
Db	1820	TGGCTTCTGGCAGTACTTGGGTCCGAGGCATCGGTGCTGTGCTCACTCTGGCCGAGT	1879
Qy	331	lGlnCysSerValSerValSerCysValArgAlaTyrGlyLysSerProSerLeuGlySe	351
Db	1880	CGAGTGCAGCGTCTCCGTCTCGTGTGTCGGGCCCTATGGGAAGTCCCCCTCCCTGGG	1939
Qy	351	rValArgAlaGlyValLeuGlyCysLeuAlaLeuAlaGlyLeuAlaAlaLeuProle	371
Db	1940	CGTTTGAGCAGGGGTCTTAGGCTGCTGGCATGGCAGGGCTGGCGCGGCTGCCCCCT	1999
Qy	371	uAlaSerValGlyGluTyrGlyAlaSerProLeuCysLeuProTyrAlaProProGluG1	391
Db	2000	GGCCTCAGTGGGAGAATACGGGGCTCCCACTCTGCTGCTGCTAGCGGCCACCTGAGGG	2059
Qy	391	GlnProAlaAlaLeuGlyPheThrValAlaLeuValMetMetAsnSerPheCysPheLe	411
Db	2060	TCAGCCAGCAGCCCTGGGCTTCCCGTGGCCCTGTGTGATGATGAATCTCTCTGTTCT	2119
Qy	411	uValValAlaGlyAlaTyrIleLysLeuTyrCysAspLeuProArgGlyAspPheGluAl	431
Db	2120	GGTGTGGCGGTGCTTACTCAAACTGTGACTGTGACCTGGCGGGGGGACTTTGAGGC	2179
Qy	431	aValTrpAspCysAlaMetValArgHisValAlaTrpLeuIlePheAlaAspGlyLeuLe	451

Search completed: April 14, 2005, 19:24:43

Job time : 665 secs

Db	2180	CGTGTGGACTGCGGCATGGTGAGGCACGTGGCTGGCTCATCTTCGCAGACGGGCTCCT	2239
Qy	451	uTyrCysProValAlaPheLeuSerPheAlaSerMetLeuGlyLeuPheProValThrPr	471
Db	2240	CTACTGTCCCGTGGCCTTCTCCTCAGCTTTGCCTCCATGCTGGGCTCTTTCCCTGTACGCC	2299
Qy	471	oGluAlaValLysSerValLeuLeuValValLeuProLeuProAlaCysLeuAsnProle	491
Db	2300	CGAGCCGCTCAAGTCTGTCTGCTGGTGTGCTGCCCCCTGCTGCTGCCTCAACCCACT	2359
Qy	491	uLeuTyrLeuLeuPheAsnProHisPheArgAspAspLeuArgArgLeuArgProArgAl	511
Db	2360	GCTGTACTGCTGCTTCAACCCCTCCCGGATGACCTTCGGCGGCTTCGGCCCCCGGC	2419
Qy	511	aGlyAspSerGlyProLeuAlaTyrAlaAlaGlyGluLeuGluLysSerSerCysAs	531
Db	2420	AGGGGACTCAGGGCCCTTAGCTATGCTGCGCGGGGAGCTGGAGAGAGAGCTCTGTGA	2479
Qy	531	pSerThrGlnAlaLeuValAlaPheSerAspValAspLeuIleLeuGluAlaSerGluAl	551
Db	2480	TTCTACCCAGGCTGGTGGTCTTCTCTGATGTGATCTCATTTCTGGAAGCTTCTGAAGC	2539
Qy	551	aGlyArgProProGlyLeuGluThrTyrGlyPheProSerValThrLeuIleSerCysG1	571
Db	2540	TGGCGGCCCTGGGCTGGAGACCTATGGCTTCCCTCAGTGACCCCTCATCTCTGTCA	2599
Qy	571	nGlnProGlyAlaProArgLeuGluGlySerHisCysValGluProGluGlyAsnHisPh	591
Db	2600	GCAGCCAGGGGCCCCAGGCTGGAGGGCAGGCATTTGTGTAGGCCAGAGGGGAACACTT	2659
Qy	591	eGlyAsnProGlnProSerMetAspGlyGluLeuLeuLeuArgAlaGluGlySerThrPr	611
Db	2660	TGGGAACCCCAACCCCTCCATGGATGGAGAACTGCTGCTGAGGGCAGAGGGATCTACGCC	2719
Qy	611	oAlaGlyGlyLeuSerGlyGlyGlyPheGlnProSerGlyLeuAlaPheAlaSe	631
Db	2720	AGCAGGTGGAGGCTTGTTCAGGGGTGGCGGCTTTTCAGCCCTCTGGCTTTGGCTTC	2779
Qy	631	rHisVal	633
Db	2780	ACACGTG	2786

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: April 14, 2005, 18:37:28 ; Search time 3621 Seconds
(without alignments)
6654.151 Million cell updates/sec

Title: US-10-664-667-5
Perfect score: 3307
Sequence: 1 NTHYRESWYACRYSRGIPG.....GGLSGGGFQPSGLAFASHV 633

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 34239544 seqs, 19032134700 residues

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+ p2n.model -DEV=xlh
-Q=/cgn2_1/USPTO spo01/US10664667/runat_13042005_184539_7277/app_query.fasta_1.775
-DB=EST -QFMT=fastap -SUFFIX=p2n.rst -MINMATCH=0.1 -LOOPEL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=biosum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR SCORE=spct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000
-USER=US10664667 @CGN 1 1 2607 @runat_13042005_184539_7277 -NCPU=6 -ICPU=3
-NO MNAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSBLOCK=100 -LONLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : EST:*
1: gb_est1:*
2: gb_est2:*
3: gb_hic:*
4: gb_est3:*
5: gb_est4:*
6: gb_est5:*
7: gb_est6:*
8: gb_gse1:*
9: gb_gse2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3140	95.0	3509	3 BC038795	BC038795 Homo sapi
2	2621	79.3	2822	3 AK052873	AK052873 Mus muscu
3	2582.5	78.1	3583	3 AK085901	AK085901 Mus muscu
4	1275	38.6	2724	9 AY411732	AY411732 Homo sapi
5	1231	37.2	2724	9 AY411734	AY411734 Mus muscu
6	1075.5	32.5	812	4 BG916782	BG916782 602816031
7	1062	32.1	819	4 BG863804	BG863804 602798701
8	1037	31.4	982	2 BF159363	BF159363 601770339
9	960	29.0	2593	9 AY411733	AY411733 Pan trogl

10	956	28.9	3035	3	AK044357	AK044357 Mus muscu
11	932.5	28.2	753	4	BG916963	BG916963 602816251
12	929	28.1	3855	3	AK040883	AK040883 Mus muscu
13	914	27.6	690	2	BB636845	BB636845 BB636845
14	887.5	26.8	869	2	BF158974	BF158974 601766476
15	803	24.3	680	7	CN258480	CN258480 170005321
16	794	24.0	640	7	CO351740	CO351740 DR_AOV_NR
17	763	23.1	3006	3	AK033699	AK033699 Mus muscu
18	690	20.9	1100	1	AL530798	AL530798 AL530798
19	646	19.5	800	4	BG196708	BG196708 RST15934
20	614	18.6	776	5	BU708439	BU708439 UI-M-FCO-
21	609	18.4	461	2	BE149368	BE149368 RCL-HT025
22	600	18.1	671	4	BM795259	BM795259 R-EST0077
23	565	17.1	880	5	BQ216343	BQ216343 AGENCOURT
24	540	16.3	798	4	BI143698	BI143698 602907345
25	535	16.2	651	2	BB621262	BB621262 BB621262
26	513.5	15.5	695	6	BY731701	BY731701 BY731701
27	507.5	15.3	788	7	CF734172	CF734172 UI-M-HBO-
28	493.5	14.9	1719	2	BF144122	BF144122 601788812
29	484	14.6	764	5	BU387687	BU387687 603582771
30	475	14.4	607	5	BQ572012	BQ572012 UI-M-FBO-
31	471	14.2	310	7	Z44115	Z44115 HSC1SG111 n
32	469.5	14.2	807	3	AK016357	AK016357 Mus muscu
33	469.5	14.2	816	6	BY716002	BY716002 BY716002
34	465	14.1	471	5	BX281802	BX281802 BX281802
35	465	14.1	771	5	BQ425502	BQ425502 AGENCOURT
36	461	13.9	839	4	BI150746	BI150746 602914834
37	439.5	13.3	2360	3	AK016635	AK016635 Mus muscu
38	434	13.1	600	4	BI989071	BI989071 4024-89 M
39	430.5	13.0	604	4	BM538630	BM538630 ha98a06.9
40	429.5	13.0	2079	9	AY399755	AY399755 Mus muscu
41	427	12.9	594	5	BX482638	BX482638 DKFZp686M
42	419	12.7	807	4	BI182677	BI182677 UNL-P-PN-
43	416	12.6	868	2	BE868472	BE868472 60144469
44	414.5	12.5	2088	9	AY399753	AY399753 Homo sapi
45	411.5	12.4	1060	1	AL541959	AL541959 AL541959

ALIGNMENTS

RESULT 1
BC038795

LOCUS
DEFINITION

BC038795 3509 bp mRNA linear HTC 19-NOV-2003
Homo sapiens leucine-rich repeat-containing G protein-coupled
receptor 6, mRNA (cdna clone IMAGE:5220507), with apparent retained
intron.

ACCESSION BC038795

VERSION BC038795.1

KEYWORDS HTC.

SOURCE Homo sapiens (human)

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE

AUTHORS

1 (bases 1 to 3509)
Strausberg, R.L., Feingold, E.A., Grouse, L.H., Derge, J.G.,
Klausner, R.D., Collins, F.S., Wagner, L., Shenmen, C.M., Schuler, G.D.,
Altschul, S.F., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Bhat, N.K.,
Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Hsieh, F.,
Diatchenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L.,
Stapleton, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L.,
Schetz, T.E., Brownstein, M.J., Ustin, T.B., Toshiyuki, S.,
Carninci, P., Prange, C., Raha, S.S., Loquellano, N.A., Peters, G.J.,
Abramson, R.D., Mullahy, S.J., Bosak, S.A., McEwan, P.J.,
McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S.,
Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hulyk, S.W.,
Villalón, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A.,
Fahey, J., Heintz, E., Kettner, M., Madan, A., Rodriguez, S.,
Sanchez, A., Whiting, M., Madan, A., Young, A.C., Shevchenko, Y.,
Bouffard, G.G., Blakeley, R.W., Touchman, J.W., Green, E.D.,
Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M.,
Butterfield, Y.S., Krzywinski, M.I., Skalek, U., Smalley, D.E.,
Schnerch, A., Schein, J.E., Jones, S.J. and Marra, M.A.
Generation and initial analysis of more than 15,000 full-length

TITLE


```
Db 2271 GCGACTTTGAGCGCTGTGGACTGCGCCCATGTGTGAGCGCAGTGGCGCTGCATCTTC 2330
Oy 447 AlahpGlyLeuLeuTyCyProValAlaPheLeuSerPheAlaSerMetLeuGlyLeu 466
Db 2331 GCAGACGGGCTCTTACTGTCCGTCGCTTCTCAGCTTCGCTCCATCTCGGCGCTC 2390
Oy 467 PheProValThrProGluAlaValLySerValLeuLeuValValLeuProLeuProAla 486
Db 2391 TTCCCTGTGACGCGCGAGCGCTCAAGTCTGTCTGTCTGGTGGTCTGCCCTGCCTGCC 2450
Oy 487 CysLeuAenProLeuLeuTyLeuLeuPheAenProHisPheArgAspLeuArg 506
Db 2451 TGCCTCAACCCACTGCTGTACTGTCTTCAACCCCTTCCGGGATGACCTTCGGCGG 2510
Oy 507 LeuArgProArgAlaGlyAspSerGlyProLeuAlaTyAlaAlaAlaGlyGluLeuGlu 526
Db 2511 CTTGGCCCCGGCGAGGACTCAGGCGCCCTAGCTATGTCTGCGCCGGGAGCTGGAG 2570
Oy 527 LysSerSerCysAspSerThrGlnAlaLeuValAlaPheSerAppValAspLeuLeu 546
Db 2571 AAGAGCTCTGTGATTCTACCCAGGCGCTGTAGCTTCTCTGTATGTGGATCTCATTTCTG 2630
Oy 547 GluAlaSerGluAlaGlyArgProGlyLeuGluThrTyGlyPheProSerValThr 566
Db 2631 GAAGCTTCTGAAGCTGGCGGCCCTCGGCTGGAGACCTATGGCTTCCCTCAGTGACC 2690
Oy 567 LeuLeuSerCysGlnProGlyAlaProArgLeuGluGlySerHisCysValGluPro 586
Db 2691 CTATCTCTGTGACGACCGAGGCCCCCGAGCTGGAGGCGAGCCATGTGTAGAGGCA 2750
Oy 597 GluGlyAenHisPheGlyAsnProGlnProSerMetAspGlyGluLeuLeuLeuArgAla 606
Db 2751 GAGGGAACCACTTTGGGAACCCCCCAACCTCCATGGATGGAGAACTGCTGCTGAGGGCA 2810
Oy 607 GluGlySerThrProAlaGlyGlyLeuSerGlyGlyGlyPheGlnProSerGly 626
Db 2811 GAGGATCTACGCCACGAGGTGGAGGCTTGTGAGGGGTGGCGGCTTTTCAGCCCTCTGGC 2870
Oy 627 LeuAlaPheAlaSerHisVal 633
Db 2871 TTGGCTTTGCTTCACACGTG 2891

RESULT 2
AK052873 2822 bp mRNA linear HTC 03-APR-2004
LOCUS Mus musculus 16 days neonate heart cDNA, RIKEN full-length enriched
DEFINITION library, clone:DB30015D13 product:CDNA FLJ14471 FIS, CLONE
MAMMA1001030, WEAKLY SIMILAR TO LUTROPIN-CHORIOGNADOTROPIC HORMONE
RECEPTOR homolog [Homo sapiens], full insert sequence.
ACCESSION AK052873
VERSION AK052873.1 GI:26095426
KEYWORDS HTC; CAP trapper.
SOURCE Mus musculus (house mouse)
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1
REFERENCE 1
AUTHORS Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,
Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
TITLE Normalization and subtraction of cap-trapper-selected cDNAs to
prepare full-length cDNA libraries for rapid discovery of new genes
JOURNAL Genome Res. 10 (10), 1617-1630 (2000)
MEDLINE 20499374
PUBMED 11042159
REFERENCE 2
AUTHORS Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,
Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
TITLE Normalization and subtraction of cap-trapper-selected cDNAs to
prepare full-length cDNA libraries for rapid discovery of new genes
JOURNAL Genome Res. 10 (10), 1617-1630 (2000)
MEDLINE 20499374
PUBMED 11042159
REFERENCE 3
AUTHORS Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P.,
Konno, H., Akiyama, J., Nishi, K., Kitsunai, T., Tashiro, H., Itoh, M.,
```

```
Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A.,
Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K.,
Fujiwaki, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watahiki, M.,
Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsuura, S., Kawai, J.,
Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.
RIKEN integrated sequence analysis (RISA) system--384-format
sequencing pipeline with 384 multicapillary sequencer
Genome Res. 10 (11), 1757-1771 (2000)
20530913
11076861
4
The RIKEN Genome Exploration Research Group Phase II Team and the
FANTOM Consortium.
Functional annotation of a full-length mouse cDNA collection
Nature 409, 685-690 (2001)
5
The FANTOM Consortium and the RIKEN Genome Exploration Research
Group Phase I & II Team.
Analysis of the mouse transcriptome based on functional annotation
of 60,770 full-length cDNAs
Nature 420, 563-573 (2002)
6 (bases 1 to 2822)
Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Bono, H., Carninci, P.,
Fukuda, S., Furuno, M., Hanagaki, T., Hara, A., Hashizume, W.,
Hayashida, K., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hirozane, T.,
Hori, F., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kasukawa, T.,
Kato, H., Kawai, J., Kojima, Y., Kondo, S., Konno, H., Kouda, M.,
Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M.,
Nakamura, M., Nishi, K., Nomura, K., Numazaki, R., Ohno, M., Oheato, N.,
Okazaki, Y., Saito, R., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N.,
Sano, H., Sasaki, D., Shibata, K., Shinagawa, A., Shiraki, T.,
Sogabe, Y., Tagami, M., Tagawa, A., Takahashi, F., Takaku-Akahira, S.,
Takeda, Y., Tanaka, T., Tomaru, A., Toya, T., Yasunishi, A.,
Muramatsu, M. and Hayashizaki, Y.
Direct Submission
Submitted (16-JUL-2001) Yoshihide Hayashizaki, The Institute of
Physical and Chemical Research (RIKEN), Laboratory for Genome
Exploration Research Group, RIKEN Genomic Sciences Center (GSC),
RIKEN Yokohama Institute, 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama,
Kanagawa 230-0045, Japan (E-mail: genome-res@gsc.riken.jp,
URL: http://genome.gsc.riken.jp/, Tel: 81-45-503-9222,
Fax: 81-45-503-9216)
cDNA library was prepared and sequenced in Mouse Genome
Encyclopedia Project of Genome Exploration Research Group in Riken
Genomic Sciences Center and Genome Science Laboratory in RIKEN.
Division of Experimental Animal Research in Riken contributed to
prepare mouse tissues.
Please visit our web site for further details.
URL: http://genome.gsc.riken.jp/
URL: http://fantom.gsc.riken.jp/.
FEATURES
Location/Qualifiers
source
1..2822
/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="FANTOM.DB:D830015D13"
/db_xref="taxon:10090"
/clone="D830015D13"
/tissue_type="heart"
/clone_lib="RIKEN full-length enriched mouse cDNA library"
/deco_stage="16 days neonate"
misc_feature
1..2822
/note="CDNA FLJ14471 FIS, CLONE MAMMA1001030, WEAKLY
SIMILAR TO LUTROPIN-CHORIOGNADOTROPIC HORMONE RECEPTOR
homolog [Homo sapiens] (SPR|Q96K69, evidence: FASTI,
85.8%ID, 92.1%length, match=637)"
ORIGIN
Alignment Scores: 2.41e-233 Length: 2822
Pred. No.: 2621.00 Matches: 518
Score: 89.12% Conservative: 31
Percent Similarity: 84.09% Mismatches: 63
Best Local Similarity:
```

```
Query Match: 79.26% Indels: 6
DB: 3 Gaps: 1
US-10-664-667-5 (1-633) x AK052873 (1-2822)

Qy 19 ProGlySerThrHisAlaSerValGluArgSerGlnGlyLeuSerLeuProAlaHisPro 38
Db 404 CTTGGAGCTGCTCATATCA-----GATCAGGAGATTATCCACCGCTGCACAGGTG 454
Qy 39 AlaSerLeuAlaAlaLeuAlaAlaSerAsnThrThrAlaSerGlyLeuGlu***Asp 58
Db 455 TCAGAGCTGGAGGAAA--TTGGCCCTCCGACATACCGGATCAAGGAAATTGGT--GCAGAT 512
Qy 59 ThrPheSerGlnLeuSerSerLeuGlnAlaLeuAspLeuSerTrpAsnAlaIleArgSer 78
Db 513 ACCTTCACCGACGTGGCTCTTGGCAAGCTTTAGACCTGAGTTGGAATGCCATCGTGCC 572
Qy 79 IleHisProGluAlaPheSerThrLeuHisSerLeuValLysLeuAspLeuThrAspAsn 98
Db 573 ATCCACCCCTGAGGCTTTCTCAACCTTCGATCTCTGGTTAAGCTGGACCTGACTGACAAC 632
Qy 99 GlnLeuThrThrLeuProLeuAlaGlyLeuGlyLeuMetHisLeuLysLeuLysGly 118
Db 633 CAGCTGACCACTGCCCTGGCTGGCTGGGAGCGCTGATGCACCTGAAGCTCAAGGG 692
Qy 119 AsnLeuAlaLeuSerGlnAlaPheSerLysAspSerPheProLysLeuArgIleLeuGlu 138
Db 693 AACTTGGCCCTGCTCAGGCCCTTCTCCAAGGACAGTTTCCCAAACTGAGGATCTCTGGAG 752
Qy 139 ValProTyrAlaTyrGlnCysCysProTyrGlyMetCysAlaSerPhePheLysAlaSer 158
Db 753 GTGCCCTACGGCTACACAGTGTGTGCTACGGCATCTGTGCAGCTTCTTCAAGACCTCT 812
Qy 159 GlyGlnTrpGluAlaGluAspLeuHisLeuAspAspGluGluSerSerLysArgProLeu 178
Db 813 GGGCAGTGGCAGCGGAGGAGCTTTCATCCAGGAGAGGAGGACCAAAAGAGGCCCTGTG 872
Qy 179 GlyLeuLeuAlaArgGlnAlaGluAsnHisTyrAspGlnAspLeuAspGluLeuGlnLeu 198
Db 873 GGTCTCTTGTGTCAGCAGCTGAGNACCCTATGACCTAGACCTGGATGAGCTCCAGATG 932
Qy 199 GluMetGluAspSerLysProHisProSerValGlnCysSerProThrProGlyProPhe 218
Db 933 GGGACAGAGGACTCAAAAGCCACACCCAGTGTCCAGTGCAGCTGAGCCCTGTGTCCAGGCCCTTC 992
Qy 219 LysProCysGluTyrLeuPheGluSerTrpGlyIleArgLeuAlaValTrpAlaIleVal 238
Db 993 AAGCCCTGCGAGCACCTCTTTGAGAGCTGGGGCATCCGCCCTTGCTGTGTGGGCCATCGTG 1052
Qy 239 LeuLeuSerValLeuCysAsnGlyLeuValLeuLeuThrValPheAlaGlyGlyProAla 258
Db 1053 TTACTCTCCGTACTCTGTAAAGGGCTGGTGTGCTGACAGTCTTTGGCAGCGAGCCAGC 1112
Qy 259 ProLeuProValLysPheValValGlyAlaIleAlaGlyAlaAsnThrLeuThrGly 278
Db 1113 CCGTGTCCTCCCGTCAAGCTTGTGGTGGTGGATGGCAGCGCGCAACGCCCTGTCTGGGC 1172
Qy 279 IleSerCysGlyLeuLeuAlaSerValAspAlaLeuThrPheGlyGlnPheSerGluTyr 298
Db 1173 ATTTCTCTGTGCTCTCTGGCCCTCGTGGACGCCCTTGACCTATGCTGCTGCTGCTGAGTAT 1232
Qy 299 GlyAlaArgTrpGluThrGlyLeuGlyCysArgAlaThrGlyPheLeuAlaValLeuGly 318
Db 1233 GGACCCCGCTGGAGAGCGGTCTGGGCTGCCAGGCTACGGCTTCTCGGCTGTCTCGGCT 1292
Qy 319 SerGluAlaSerValLeuLeuLeuThrLeuAlaAlaValGlnCysSerValSerValSer 338
Db 1293 TCAGAGCGGTGGGTGCTGCTCTCACACTGTCGCGCGCTGGTGCAGTGCAGCATCTCGGTGACC 1352
Qy 339 CysValArgAlaTyrGlyLysSerProSerLeuGlySerValArgAlaGlyValLeuGly 358
Db 1353 TGCCTCCAGGCTACGGGAGGCGCGCTGCGCTGCGCAGCGTCCGCGAGCGGCACTGGGA 1412
```

```
Qy 359 CysLeuAlaLeuAlaGlyLeuAlaAlaAlaLeuProLeuAlaSerValGlyGluTyrGly 378
Db 1413 TGCCTGGCGCTGGCGGGCTGGCGCAGCACTGCCGCTGGCTCGTGGGAGAGATATGGC 1472
Qy 379 AlaSerProLeuCysLeuProTyrAlaProProGluGlyGlnProAlaAlaLeuGlyPhe 398
Db 1473 GCCTTCCCACACTCTGCTTGCCTACGCCCCACCCAGGCGCGCGGCTGCCCTGGGCTTC 1532
Qy 399 ThrValAlaLeuValMetMetAsnSerPheCysPheLeuValValAlaGlyAlaTyrIle 418
Db 1533 GCTGTAGCCCTGGGATGATGAACCTCGCTTCTTCTGGTGGTGGCGGCGCTCATATC 1592
Qy 419 LysLeuTyrCysAspLeuProArgGlyAspPheGluAlaValTrpAspCysAlaMetVal 438
Db 1593 AAGCTCTACTGTGACCTGCCACGGGGTGACTTTGAGCGCGTGTGGGACTGGCCATGGTG 1652
Qy 439 ArgHisVal-AlaTrpLeuIlePheAlaAspGlyLeuLeuTyrCysProValAlaPheLe 458
Db 1653 CGCCACCTGGCGCTGGCTCATCTTTGCAGATGGCTCTCTACTGCCCGCTGGCTTCTCT 1712
Qy 458 uSerPheAlaSerMetLeuGlyLeuPheProValThrProGluAlaValLysSerValLe 478
Db 1713 CAGCTTTGCTCTATGCTGGGCTCTTCCCTGTACCCCCGAGGCTGTCAAGTCAGTCTCT 1772
Qy 478 uLeuValValLeuProLeuProAlaCysLeuAsnProLeuLeuTyrLeuLeuPheAsnPr 498
Db 1773 TCTGGTGGTCTGCCCTGCTGCTGCCCTCAACCCACTGCTCTACTGCTCTTCAACCC 1832
Qy 498 oHisPheArgAspAspLeuArgArgLeuArgProArgAlaGlyAspSerGlyProLeuAl 518
Db 1833 TCACCTTCGGGATGACCTTCGGGGCTCTGGCCAAAGCCCTCGGTCGCCAGGCGCCCTAGC 1892
Qy 518 aTyrAlaAlaAlaGlyLeuLeuGlyLysSerCysAspSerThrGlnAlaLeuValAl 538
Db 1893 CTACGCTGCAGCGCGTGGTGGAGAGAGCTCTGCGACTCTCCACCACCAAGCGCTGGTGGC 1952
Qy 538 aPheSerAspValAspLeuIleLeuGluAlaSerGluAlaGlyArgProGlyLeuGly 558
Db 1953 TTTCTCAGATGTGAATCTTATCTGGNAGCTTCTGAGGCTGGGAGGCTCTCTGGGCTAGA 2012
Qy 558 uThrTyrGlyPheProSerValThrLeuIleSerCysGlnGlnProGlyAlaProArgLe 578
Db 2013 GACCTATGGCTTCCCTTCAGTGACCTCATCTCCGACATCAGCGGGGCTACCAAGCT 2072
Qy 578 uGluGlySerHisCysValGluProGluGlyAsnHisPheGlyAsnProGlnProSerMe 598
Db 2073 GGAGGGAAACCATTTGTAGATCTGATGGAAACCAAGTTTGGGAACCCACCACTCCCAT 2132
Qy 598 tAspGlyGluLeuLeuArgAlaGluGlySerThrProAlaGlyGlyGlyLeuSerGly 618
Db 2133 GAAGGGAGAACTGCTGCTGNAGCAGAGGAGGACATTTGGCAGGCTGTGGCTTCCGT 2192
Qy 618 yGlyGlyGlyPheGlnProSerGlyLeuAlaPheAlaSerHisVal 633
Db 2193 GGGTGGAGCCCTCTGCGCCCTCTGGCTCTCTCTTTGGCTCTCACTGTG 2238
```

RESULT 3

AK085901

LOCUS

DEFINITION

Mus musculus 16 days neonate heart cDNA, RIKEN full-length enriched library, clone:D830026M09 product:CDNA FJ14471 FIS, CLONE MAMMAL001030, WEAKLY SIMILAR TO LUTROPIN-CHORIOGNADOTROPIC HORMONE RECEPTOR homolog (Homo sapiens), full insert sequence.

AK085901

VERSION AK085901.1 GI:26103061

KEYWORDS HTC; CAP trapper.

SOURCE Mus musculus (house mouse)

ORGANISM

Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.

REFERENCE

1 Carninci, P. and Hayashizaki, Y.

High-efficiency full-length cDNA cloning

```

JOURNAL Meth. Enzymol. 303, 19-44 (1999)
MEDLINE 99279253
PUBMED 10349636
REFERENCE 2
AUTHORS Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,
Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M., and Hayashizaki, Y.
TITLE Normalization and subtraction of cap-trapper-selected cDNAs to
prepare full-length cDNA libraries for rapid discovery of new genes
JOURNAL Genome Res. 10 (10), 1617-1630 (2000)
MEDLINE 20499374
PUBMED 11042159
REFERENCE 3
AUTHORS Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P.,
Konno, H., Akiyama, J., Nishi, K., Kitsunai, T., Tashiro, H., Itoh, M.,
Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A.,
Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K.,
Fujiwaka, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watahiki, M.,
Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsuura, S., Kawai, J.,
Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A., and Hayashizaki, Y.
TITLE RIKEN integrated sequence analysis (RISA) system--384-format
sequencing pipeline with 384 multicapillary sequencer
JOURNAL Genome Res. 10 (11), 1757-1771 (2000)
MEDLINE 20530913
PUBMED 11076861
REFERENCE 4
AUTHORS The RIKEN Genome Exploration Research Group Phase II Team and the
FANTOM Consortium.
TITLE Functional annotation of a full-length mouse cDNA collection
JOURNAL Nature 409, 685-690 (2001)
REFERENCE 5
AUTHORS The FANTOM Consortium and the RIKEN Genome Exploration Research
Group Phase I & II Team.
TITLE Analysis of the mouse transcriptome based on functional annotation
of 60,770 full-length cDNAs
JOURNAL Nature 420, 563-573 (2002)
REFERENCE 6
AUTHORS Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Bono, H., Carninci, P.,
Fukuda, S., Furuno, M., Hanagaki, T., Hara, A., Hashizume, W.,
Hayashida, K., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hirozane, T.,
Hori, F., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kasukawa, T.,
Kato, H., Kawai, J., Kojima, Y., Kondo, S., Konno, H., Kouda, M.,
Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M.,
Nakamura, M., Nishi, K., Nomura, K., Numazaki, R., Ohno, M., Ohsato, N.,
Okazaki, Y., Saito, R., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N.,
Sano, H., Sasaki, D., Shibata, K., Shinagawa, A., Shiraki, T.,
Sogabe, Y., Tagami, M., Tagawa, A., Takahashi, F., Takaku-Akahira, S.,
Takeda, Y., Tanaka, T., Tomaru, A., Toya, T., Yasunishi, A.,
Muramatsu, M., and Hayashizaki, Y.
TITLE Direct Submission
JOURNAL Submitted (16-APR-2002) Yoshihide Hayashizaki, The Institute of
Physical and Chemical Research (RIKEN), Laboratory for Genome
Exploration Research Group, RIKEN Genomic Sciences Center (GSC),
RIKEN Yokohama Institute, 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama,
Kanagawa 230-0045, Japan (E-mail: genome-res@gsc.riken.jp,
URL: http://genome.gsc.riken.jp/, Tel: 81-45-503-9222,
Fax: 81-45-503-9216)
COMMENT cDNA library was prepared and sequenced in Mouse Genome
Encyclopedia Project of Genome Exploration Research Group in Riken
Genomic Sciences Center and Genome Science Laboratory in RIKEN.
Division of Experimental Animal Research in Riken contributed to
prepare mouse tissues.
Please visit our web site for further details.
URL: http://genome.gsc.riken.jp/
URL: http://fantom.gsc.riken.jp/.
FEATURES source
location/Qualifiers
1. .3583
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="FANTOM DB:D830026M09"
/db_xref="caxon:10090"
/clone="D830026M09"
/tissue_type="heart"
misc_feature
/clone_lib="RIKEN full-length enriched mouse cDNA library"
/dev_stage="16 days neonate"
95. .3002
/note="CDNA FLJ14471 FIS, CLONE MAMMA1001030, WEAKLY
SIMILAR TO LUTROPIN-CHORIOGONADOTROPIC HORMONE RECEPTOR
homolog [Homo sapiens] (SPR|Q96K69, evidence: PASTY,
85.8%id, 92.1%length, match=63)"
putative"
3563. .3568
/note="putative"
3583
/note="putative"
polyA_site
3583
polyA_site
3583
ORIGIN
Alignment Scores:
Pred. No.: 1,39e-229 Length: 3583
Score: 2582.50 Matches: 514
Percent Similarity: 88.19% Conservative: 31
Best Local Similarity: 83.17% Mismatches: 67
Query Match: 78.09% Indels: 7
DB: 3 Gaps: 1
US-10-664-667-5 (1-633) x AK085901 (1-3583)
QY 19 ProGlySerThrHisAlaSerValGluArgSerGlnGlyLeuSerLeuProAlaHisPro 38
DB 1162 CCTGGAGCTGCTCATATCA-----GATCAGGAGTATCCAGCCTGGCAGGT 1212
QY 39 AlaSerLeuAlaLeuAlaAlaSerAsnThrThrAlaSerGlyLeuGlu***Aap 58
DB 1213 GTCAGAGCTGGAGGAATTTGGCTCCGACATACCGGATCAAGGAATTTGT-CCAGAT 1271
QY 59 ThrPheSerGlnLeuSerSerLeuGlnAlaLeuAspLeuSerTrpAsnAlaIleArgSer 78
DB 1272 ACCTTCAGCCAGCTGGGCTCCTTCGACAGCTTTAGACCTGAGTTGGAATGCCATCGTGCC 1331
QY 79 IleHisProGluAlaPheSerThrLeuHisSerLeuValLeuAspLeuThrAspAen 98
DB 1332 ATCCACCCTGAGGCTTTCTCAACCTTCGATCTTGGTTAAGCTGGACCTGACTGACAAAC 1391
QY 99 GlnLeuThrThrLeuProLeuAlaGlyLeuGlyLeuMetHisLeuLeuLeuGly 118
DB 1392 CAGCTGACCAACACATGCCCCCTGGCTGGGCTGGAGGCTGTGACCTGAACTCAAGGG 1451
QY 119 AsnLeuAlaLeuSerGlnAlaPheSerTrpAspSerPheProLeuLeuArgIleLeuGlu 138
DB 1452 AACTTGGCCCTGCTCAGGCCCTTCTCCAGAGACAGTTTCCAAAACCTGAGGATCCTGGAG 1511
QY 139 ValProTyrAlaTyrGlnCysCysProTyrGlyMetCysAlaSerPhePheLeuAlaSer 158
DB 1512 GTGCCCTACGCTACCACTGCTGTGCCTACGGCATCTGTGCAGCTTCTTCAAGACCTCT 1571
QY 159 GlyGlnTrpGluAlaGluAspLeuHisLeuAspAspGluGluSerSerLeuArgProLeu 178
DB 1572 GGGCAGTGGCAGCGCAGGACTTTTCATCCAGAGAGAGGAGGCCCAAAAGAGGCCCTG 1631
QY 179 GlyLeuLeuAlaArgGlnAlaGluAsnHisTrpAspGlnAspLeuAspGluLeuGlnLeu 198
DB 1632 GGTCTCTTCTGGACACAGCTGAGAACCATATGACCTAGACCTGGATGAGCTCCAGATG 1691
QY 199 GluMetGluAspSerLeuProHisProSerValGlnCysSerProThrProGlyProPhe 218
DB 1692 GGGACAGAGGACTCAAGGCCACACCCAGTGTCCAGTGCAGCCCTGTTCCAGGCCCTTC 1751
QY 219 Lys-ProCysGluTyrLeuPheGluSerTrpGlyIleArgLeuAlaValTrpAlaIleVa 238
DB 1752 AAGGCCCTGCGAGCACCTCTTTTGAGAGCTGGGGCATCCGCTTGTGTGGGCCATCGT 1811
QY 238 lLeuLeuSerValLeuCysAsnGlyLeuValLeuLeuThrValPheAlaGlyGlyProAl 258
DB 1812 GTTACTCTCCGTACTCTGTAAAGGGCTGGTCTGCTGACAGTCTTTTGCAGGAGGCCAG 1871
QY 258 aProLeuProValLysPheValValGlyAlaIleAlaGlyAlaAsnThrLeuThrGl 278

```

```
Db 1872 CCGCGTGTCCCGCTCAAGCTTGTGGTGGATGGCAGCGCCCAACGCCCTGTGCGG 1931
Qy 278 yIleSerCysGlyLeuLeuAlaSerValAspAlaLeuThrPheGlyGlnPheSerGluTy 298
Db 1932 CATTTCTGCTGCTCTCCGGCTGGTGGACGCCCTTGACCTATGCTAGTTGCTGAGTA 1991
Qy 298 rGlyAlaArgTrpGluThrGlyLeuGlyCysArgAlaThrGlyPheLeuAlaValLeuG1 318
Db 1992 TGGAGCCCGCTGGAGAGCGGTCTGGCTGCGCAGGTACGGGCTCTCCCTGGCTGTCTCGG 2051
Qy 318 ySerGluAlaSerValLeuLeuLeuThrLeuAlaAlaValGlnCysSerValSerValSe 338
Db 2052 TTCAGAGCGCTGGTGTGCTGCTCACACTGGCGCGCTGCGAGTCAGCATCTCGGTGAC 2111
Qy 338 rCysValArgAlaTyTrpGlyLysSerPro-SerLeuGlySerValArgAlaGlyValLeuG 358
Db 2112 CTGCGTCCGAGCTTACGGGAAGCGCGCTGCGCTGGCAGCGTCTCGCGCAGCGCACCTTG 2171
Qy 358 lyCysLeuAlaLeuAlaGlyLeuAlaAlaLeuProLeuAlaSerValGlyGluTyTrG 378
Db 2172 GATGCTCGCTGGCTGGCGGGCTGGCGCGCAGCACTGCCCTGGCTGGTGGAGAGTATG 2231
Qy 378 lyAlaSerProLeuCysLeuProTyAlaProProGluGlyGlnProAla-AlaLeuGly 397
Db 2232 GCGCTCCCGCACTCTGCTGCTAGCCCTACGCCCCACCGCGCGCGCTGCCCTGGCGG 2291
Qy 398 PheThrValAlaLeuValMetMetAsnSerPheCysPheLeuValValAlaGlyAlaTy 417
Db 2292 TTGCGTGTAGCCCTGGTGATGATGAACCTCGCTCTGCTTCTGGTGGTGGCGGCGCTAC 2351
Qy 418 lIeIysLeuTyTrpCysAspLeuProArgGlyAspPheGluAlaValTrpAspCysAlaMet 437
Db 2352 ATCAAGCTCTACTGTGACCTGCCCGGGTGACTTTGAGGGCGGTGGGAGCTGGCGCCATG 2411
Qy 438 ValArgHisValAlaTrpLeuIlePheAlaAspGlyLeuLeuTyTrpCysProValAlaPhe 457
Db 2412 GTGCGCCACGTGGCTGTGCTCATCTTTCAGAGTGGCTCTCTACTGCGCGCTGGCTTC 2471
Qy 458 leuSerPheAlaSerMetLeuGlyLeuPheProValThrProGluAlaValLysSerVal 477
Db 2472 CTCAGCTTTGCTCTATGCTGGGCTCTTTCGCTGTGACCCCGCGAGGCTGTCAAGTCA 2531
Qy 478 leuLeuValValLeuProLeuProAlaCysLeuAsnProLeuLeuTyTrpLeuPheAsn 497
Db 2532 CTTCTGGTGGTGTGCTGCTGCTGCTGCTGCTCAACCCACCTGCTCTACTGCTCTTAAAC 2591
Qy 498 ProHisPheArgAspAspLeuArgArgLeuArgProArgAlaGlyAspSerGlyProLeu 517
Db 2592 CTTCACTTCGGGATGACCTTCGGGGCTCTGGCCAGCCCTCGGTCCCGAGGGCCCTA 2651
Qy 518 AlaTyAlaAlaAlaGlyGluLeuGlyLysSerSerCysAspSerThrGlnAlaLeuVal 537
Db 2652 GCCTACGCTGAGCGCGGTGAGCTGAGAGAGCTCTCGAGCTCCGACTCAAGCGCTGTG 2711
Qy 538 AlaPheSerAspValAspLeuIleLeuGluAlaSerGluAlaGlyArgProProGlyLeu 557
Db 2712 GCTTTCTCACATGTGGATCTTATTCTGGAAGCTTCTGAGGCTGGCGAGCCCTCTGGGGTA 2771
Qy 558 GluThrTyTrpGlyPheProSerValThrLeuIleSerCysGlnGlnProGlyValaProArg 577
Db 2772 GAGACCTATGCTTCCCTTCAGTGACCTCATCTCCGACATCAGCCGGGGCTACCCAG 2831
Qy 578 leuGluGlySerHisCysValGluProGluGlyAsnHisPheGlyAsnProGlnProSer 597
Db 2832 CTGGAGGGAACCACTTTGTAGAGTCTGATGGAACCAAGTTTGGGAACCCCAACCTCCC 2891
Qy 598 MetAspGlyGluLeuLeuLeuArgAlaGlyLeuSerThrProAlaGlyGlyLeuSer 617
Db 2892 ATGAAGGGAACCTGCTCTGAAGGAGGAGGACCTTTGGCAGGCTGTGGCTCTTCC 2951
Qy 618 GlyGlyGlyGlyPheGlnProSerGlyLeuAlaPheAlaSerHisVal 633
```

```
Db 2952 GTGGGTGGAGCCCTCTGGCCCTCTGGCTCTCTCTTTGGCTCTCACTTG 2999
RESULT 4
AY411732 2724 bp DNA linear GSS 16-DEC-2003
LOCUS Homo sapiens GPR49 gene, VIRTUAL TRANSCRIPT, partial sequence,
DEFINITION genomic survey sequence.
ACCESSION AY411732
VERSION AY411732.1 GI:39767700
KEYWORDS GSS.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 2724)
AUTHORS Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A.,
Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
Ferreira,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
Adams,M.D. and Cargill,M.
TITLE Inferring nonneutral evolution from human-chimp-mouse orthologous
gene trios
JOURNAL Science 302 (5652), 1960-1963 (2003)
PUBMED 14671302
REFERENCE 2 (bases 1 to 2724)
AUTHORS Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A.,
Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
Ferreira,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
Adams,M.D. and Cargill,M.
TITLE Direct Submission
JOURNAL Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,
Rockville, MD 20850, USA
COMMENT This sequence was made by sequencing genomic exons and ordering
them based on alignment.
FEATURES
source
location/Qualifiers
1..2724
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
<1..>2724
/gene="GPR49"
/locus_tag="HCM4322"
gene
ORIGIN
Alignment Scores:
Pred. No.: 1,51e-107 Length: 2724
Score: 1275.00 Matches: 250
Percent Similarity: 64.92% Conservative: 96
Best Local Similarity: 46.90% Mismatches: 157
Query Match: 38.55% Indels: 30
DB: Gaps: 8
US-10-664-667-5 (1-633) x AY411732 (1-2724)
Qy 54 LysLeuGluV**AapThrPheSerGlnLeuSerSerLeuGlnAlaLeuAspLeuSerTrp 73
Db 1162 GAAATTAAGATTGACATTTCCAGCAGTTCCTTAGCTCCGATCGCTGAATTTGGCTGG 1221
Qy 74 AsnAlaIleArgSerIleHisProGluAlaPheSerThrLeuHisSerLeuValLysLeu 93
Db 1222 AACAAAAATGCTATTATTATCCACCAATGCATTTTCCACTTTGCCATCCCTAATAAGCTG 1281
Qy 94 AspLeuThrAspAsnGlnLeuThrThrLeuProLeuAlaGlyLeuGlyLeuMetHis 113
Db 1282 GACCTATCGTCCACACCTCCCTGTGCTCTTTCTATAAATCGGTTACATGGTTTAACTCAC 1341
Qy 114 LeuLysLeuLysGlyAsnLeuAlaLeuSerGlnAlaPheSerLysAspSerPheProLys 133
Db 1342 TTAAATTAACAGGAATCATGCTTACAGAGCTTGATATCATCTGAAACTTTCCAGAA 1401
Qy 134 LeuArgIleLeuGluValProTyAlaTyTrpGlnCysCysProTyTrpGlyMetCysAlaSer 153
Db 1402 CTCAAGGTTATAGAAATGCTTATGCTTACAGTGTGTGATTTGGAGTGTGTGAGAA 1461
```


Alignment Scores:
Pred. No.: 1.03e-89 Length: 812
Score: 1075.50 Matches: 223
Percent Similarity: 88.41% Conservative: 21
Best Local Similarity: 80.80% Mismatches: 27
Query Match: 32.52% Indels: 8
DB: 4 Gaps: 3

US-10-664-667-5 (1-633) x BG916782 (1-812)

Qy 131 PheProLysLeuArgIleLeuGluValProTyrAlaTyrGlnCysCysProTyrGlyMet 150
Db 1 TTTCCAAAACAGGAGATCTGGAGTGGCTACGCTACCAAGTGTGCTGCTACGGCATC 60

Qy 151 CysAlaSerPhePheLysAlaSerGlyGlnTyrGluAlaGluAspLeuHisLeuAspAsp 170
Db 61 TGTGCCAGCTTCTTCAAGACCTCTGGCAGTGGCGAGGCGGAGGATTTTCATCCAGAGAA 120

Qy 171 GluGluSerSerLysArgProLeuGlyLeuLeuAlaArgGlnAlaGluAsnHisTyrAsp 190
Db 121 GAGGAGGCCAACAGAGGCCCCCTGGTCTCTCTTGTGGACAGCTGAGAACCCACTATGAC 180

Qy 191 GlnAspLeuAspGluGlnLeuGluMetGluAspSerLysProHisProSerValGln 210
Db 181 CTAGACCTGGATGAGCTCCAGATGGGACAGAGGACTCAAGGCCAAACCCAGTGTCCAG 240

Qy 211 CysSerProThrProGlyProPheLysProCysGluTyrLeuPheGluSerTrpGlyIle 230
Db 241 TGCAGCCCTGTTCAGAGCCCTTCAAGCCCTGCGAGCACCTCTTTGAGAGCTGGGGCATC 300

Qy 231 ArgLeuAlaValTrpAlaIleValLeuLeuSerValLeuCysAsnGlyLeuValLeuLeu 250
Db 301 CGCCTTGTGTGGGCCATCGTGTACTCTCCGTACTCTGTAAAGGGCTGGTGTGTG 360

Qy 251 ThrValPheAlaGlyGlyProAlaProLeuProProValLysPheValValGlyAlaIle 270
Db 361 ACAGTCTTTGCCAGCGGACAGC-CCGCTGTCCCCCGTCAAGCTTGTGGTGGTGGCATG 419

Qy 271 AlaGlyAlaAsnThrLeuThrGlyLeuSerCysGlyLeuLeuAlaSerValAspAlaLeu 290
Db 420 GCAGCGGCCAACGCCCTTACGGGGCATTTCTGTGTCTCTGGCTCTGGTGGACCCCTG 479

Qy 291 ThrPheGlyGlnPheSerGluTyrGlyAlaArgTrpGluThrGlyLeuGlyCysArgAla 310
Db 480 ACCTATGTCTGCTGAGTATGGAGCCCGCTGGAGAGCGGTCTGGGCTGCCAGCT 539

Qy 311 ThrGlyPheLeuAlaValLeuGlySerGluAlaSerValLeuLeuLeuThrLeuAlaAla 330
Db 540 ACGGCTTCTTGGCTGTCTGGTTCAGAGGGCTGGTGTCTGTCTGTCTCACACTGGCGCG 599

Qy 331 ValGlnCysSerValSerValSerCysValArgAlaTyrGlyLysSerProSerLeuGly 350
Db 600 GTGCAGTGCAGCATCTGTGACCTGCTGCGTCCAGCTTACGGGAAGCGCCGCTGCT-GGC 658

Qy 351 SerValArgAlaGlyValLeuGlyCysValLeuAlaLeuAlaGlyLeuAlaAlaLeuPro 370
Db 659 ACGTCCGCGCAGGCGCATGGGATGCTTGGCGCTGGCGGCTGGCGGCTGGCGAGCACTGC 715

Qy 371 LeuAlaSerValGlyGluTyrGlyAlaSerProLeuCysLeuProTyrAlaProGlu 390
Db 716 CGTGGCTGGTGGAGATGTCGCTC-CCAATTCTGTCTGCT-AGCCACCG- 767

Qy 391 GlyGln-ProAlaAlaLeuGlyPheThrValAlaLeuValMetMet 405
Db 768 ---AGGGCGGGCGGCTGGTGTGTGTATCCCTTGGTGGTGTGATG 810

RESULT 7
BG863804
LOCUS 602798701F1 NCI_CGAP_Mam4 Mus musculus cDNA clone IMAGE:4934280 5',
DEFINITION mRNA sequence.
ACCESSION BG863804

US-10-664-667-5 (1-633) x BG863804 (1-819)

Qy 46 AlaSerAsnThrThrAlaSerGlyLysLeuGlu**AspThrPheSerGlnLeuSerSer 65
Db 42 GCCTCCGACATACAGGATCAAGGAATGTGT-GCAGATACCTTCAGCAGCTGGGCTCC 100

Qy 66 LeuGlnAlaLeuAspLeuSerTrpAsnAlaIleArgSerIleHisProGluAlaPheSer 85
Db 101 TTGCAGCTTTAGACCTGAGTTGGATGCCATCCGTGGCCATCCACCTGAGGCTTTCTCA 160

Qy 86 ThrLeuHisSerLeuValLysLeuAspLeuThrAspAsnGlnLeuThrThrLeuProLeu 105
Db 161 ACCCTTCGATCCTTGGTTAAGCTGGAGCTGACTGACAAACAGCTGACCACTGCCCTG 220

Qy 106 AlaGlyLeuGlyGlyLeuMetHisLeuLysLeuGlyAsnLeuAlaLeuSerGlnAla 125
Db 221 GCTGGCTGGGAGGCGCTGATCCTGAAGCTCAAGGGAACCTGGCCCTGTCTCAGGCC 280

Qy 126 PheSerLysAspSerPheProLysLeuArgIleLeuGluVal-ProTyrAlaTyrGlnCys 145
Db 281 TTCTCCAAAGGACAGTTTCCCAAACTGAGGATCTCTGGAGGTGCCCTACGCTACAGTG 340

Qy 145 sCysProTyrGlyMetCysAlaSerPhePheLysAlaSerGlyGlnTrpGluAlaGluAs 165

Alignment Scores:
Pred. No.: 1.91e-88 Length: 819
Score: 1062.00 Matches: 215
Percent Similarity: 87.74% Conservative: 14
Best Local Similarity: 82.38% Mismatches: 29
Query Match: 32.11% Indels: 4
DB: 4 Gaps: 0

US-10-664-667-5 (1-633) x BG863804 (1-819)

Qy 46 AlaSerAsnThrThrAlaSerGlyLysLeuGlu**AspThrPheSerGlnLeuSerSer 65
Db 42 GCCTCCGACATACAGGATCAAGGAATGTGT-GCAGATACCTTCAGCAGCTGGGCTCC 100

Qy 66 LeuGlnAlaLeuAspLeuSerTrpAsnAlaIleArgSerIleHisProGluAlaPheSer 85
Db 101 TTGCAGCTTTAGACCTGAGTTGGATGCCATCCGTGGCCATCCACCTGAGGCTTTCTCA 160

Qy 86 ThrLeuHisSerLeuValLysLeuAspLeuThrAspAsnGlnLeuThrThrLeuProLeu 105
Db 161 ACCCTTCGATCCTTGGTTAAGCTGGAGCTGACTGACAAACAGCTGACCACTGCCCTG 220

Qy 106 AlaGlyLeuGlyGlyLeuMetHisLeuLysLeuGlyAsnLeuAlaLeuSerGlnAla 125
Db 221 GCTGGCTGGGAGGCGCTGATCCTGAAGCTCAAGGGAACCTGGCCCTGTCTCAGGCC 280

Qy 126 PheSerLysAspSerPheProLysLeuArgIleLeuGluVal-ProTyrAlaTyrGlnCys 145
Db 281 TTCTCCAAAGGACAGTTTCCCAAACTGAGGATCTCTGGAGGTGCCCTACGCTACAGTG 340

Qy 145 sCysProTyrGlyMetCysAlaSerPhePheLysAlaSerGlyGlnTrpGluAlaGluAs 165

VERSION BG863804.1 GI:14214342
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
TITLE NIH-MGC http://mgc.nci.nih.gov/.
JOURNAL National Institutes of Health, Mammalian Gene Collection (MGC)
COMMENT Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapb-remail.nih.gov
Tissue Procurement: Lothar Hennighausen Ph.D., Priscilla Furth
Ph.D.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM10863 row: a column: 01
High quality sequence stop: 758.
Location/Qualifiers
1. .819
/organism="Mus musculus"
/mol_type="mRNA"
/strain="NMRI"
/db_xref="taxon:10090"
/clone="IMAGE:4934280"
/tissue_type="tumor, gross tissue"
/dev_stage="5 months"
/lab_host="DH10B"
/clone_lib="NCI CGAP Mam4"
/note="Foran: mammary; Vector: pCMV-SPORT6; Site 1: SalI;
Site 2: NotI; Cloned unidirectionally. Primer: Oligo dT.
Library constructed by Life Technologies. Investigators
providing samples: Lothar Hennighausen/Priscilla Furth,
NIH Reference for transgenic model: Li et al., Cell Growth
and Differentiation 7, 3-11 (1996)."

Db 341 CTGTGCTTACGGCATCTGTGCCAGCTTCTTCAAGACCTCTCTGGCAGTGGCAGGCGCGAGA 400
Qy 165 pLeuHisLeuAspAspGluGluSerSerLeuArgProLeuGlyLeuLeuAlaAArgGlnAl 185
Db 401 CTTTCATCCAGAAAGAGGAGGACCAAGAGGCGCCCTGGGTCTCTCTTGTGAGCAAGC 460
Qy 185 aGluAsnHisTyRAspGlnAspLeuAspGluLeuGlnLeuMetGluAspSerLeuPr 205
Db 461 TGAGAACCATATGACCTAGACCTGGATGAGCTCCAGATGGGACAGAGGACTCAAGCC 520
Qy 205 oHisProSerValGlnCysSerProThrProGlyProPhelYsProCysGluTyRLeuPh 225
Db 521 AAACCCAGTGTCCAGTGCAGCCCTGTTCAGGCGCCCTTCAAGCCCTGCGAGCACCTCTT 580
Qy 225 eGluSerTrpGlyLeuArgLeuAlaValTrpAlaLeuValLeuLeuSerValLeuCysAs 245
Db 581 TGAGAGCTGGGGCATCCGCCCTGTGTGTGGCCATCGTGTACTCTCCGACTCTGTAA 640
Qy 245 nGlyLeuValLeuLeuThrValPheAlaGlyGlyProAlaProLeuProProVallyAsPh 265
Db 641 CGGCTGT 700
Qy 265 eValValGlyAlaLeuAlaGlyAlaAsnThrLeuThrGlyLeuSerCysGlyLeuLeuAl 285
Db 701 TGTGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 760
Qy 285 aSerValAspAlaLeuThrPhe-GlyGlnPheSerGlu-TyrGlyAlaAArgTrpGlu 303
Db 761 CTCGGTGCAGGCCTTGACCTATGGGTGAGGTGCGGTGAGGTATTCGGCGCGTGGGAC 817

RESULT 8

BF159363 601770339P1 NCI_CGAP_Lu29 Mus musculus cDNA clone IMAGE:3989576 5',
LOCUS mRNA sequence.

BF159363.1 GI:11039462

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

Mus musculus (house mouse)
Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 982)
NIH-MGC <http://mgc.nci.nih.gov/>.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: Gilbert Smith, Ph.D.
cDNA Library Preparation: Life Technologies, Inc.
DNA Sequencing by: The I.M.A.G.E. Consortium (LLNL)
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
Plate: LHAM9199 row: n column: 09
High quality sequence stop: 663.
Location/Qualifiers

FEATURES

source

1. 982
/organism="Mus musculus"
/mol_type="mRNA"
/strain="CZECH II"
/db_xref="taxon:10090"
/clone="IMAGE:3989576"
/tissue_type="spontaneous tumor, metastatic to mammary."
stem cell origin."
/lab_host="DH10B"
/clone_lib="NCI_CGAP_Lu29"
/note="Organ: lung; Vector: pCMV-SPORT6; Site 1: SalI;
Site 2: NotI; Cloned unidirectionally. Primer: Oligo dt.
Library constructed by Life Technologies. Investigator
providing samples: Gilbert Smith, NIH"

ORIGIN

Alignment Scores: 5.5e-86 Length: 982
Score: 1037.00 Matches: 209
Percent Similarity: 87.60% Conservativeness: 10
Best Local Similarity: 83.60% Mismatches: 31
Query Match: 31.36% Indels: 2
DB: 2 Gaps: 0
US-10-664-667-5 (1-633) x BF159363 (1-982)
Qy 378 GlyAlaSerProLeuCysLeuProTyRAlaProGluGlyGlnProAlaLeuGly 397
Db 1 GGCSCCTCCCACTCTCGCTGCCCTACGCCCCACCGAGGGCGCGCCCTGGGC 60
Qy 398 PheThrValAlaLeuValMetMetAsnSerPheCysPheLeuValValAlaGlyAlaTyR 417
Db 61 TTGCTGTAGCCCTGGGTGATGAACTCGCTCTCTGCTTCCTGGTGGTGGCGGCTTAC 120
Qy 418 IleIysLeuTyRLeuValMetMetAsnSerPheCysPheLeuValValAlaGlyAlaMet 437
Db 121 ATCAAGCTCTACTGTGACCTGCCACGGGGTGACTTTGAGGCCGTGGGACTGGCCATG 180
Qy 438 ValArgHisValAlaTrpLeuIlePheAlaAspGlyLeuLeuTyRLeuValAlaPhe 457
Db 181 GTGGCCACCTGGCTGGCTGCTCATCTTTGCAGATGGCTCTCTACTGCCCGTGGCTTC 240
Qy 458 LeuSerPheAlaSerMetLeuGlyLeuPheProValThrProGluAlaValLysSerVal 477
Db 241 CTCAGCTTTGCCCTATGCTGGGCTCTTCCCTGTCAACCCCGAGGCTGTCAAGTCAGTC 300
Qy 478 LeuLeuValValLeuProLeuProAlaCysLeuAsnProLeuLeuTyRLeuLeuPheAsn 497
Db 301 CTTCTGT 360
Qy 498 ProHisPheArgAspAspLeuArgArgLeuArgProArgAlaGlyAspSerGlyProLeu 517
Db 361 CCTCCTTCGGGATGACCTTCGGGGCTCTGGCCAGCCCTAGTCCCGGCGCCCTA 420
Qy 518 AlaTyRAlaAlaGlyGluLeuGluLysSerSerCysAspSerThrGlnAlaLeuVal 537
Db 421 GCCTACGCTCAGCGGTGAGCTGGAGAGAGGTCTCTGCGACTCCAGCAAGCGCTGGTG 480
Qy 538 AlaPheSerAspValAspLeuIleLeuGluAlaSerGluAlaGlyArgProGlyLeu 557
Db 481 GCCTTCTCAGATGTGGATCTTATCTGGAAGCTTCTGAGGCTGGGCGAGCTCTCTGGGCTA 540
Qy 558 GluThrTyRLeuPheProSerValThrLeuIleSerCysGlnGlnProGlyAlaProArg 577
Db 541 GAGACCTATGGCTTCCCTTCAGTGGACCTCATCTCCGACATCAGCCGGGGGCCACCGAG 600
Qy 578 LeuGluGlySerHisCysValGluProGluGlyAsnHisPheGlyAsnProGlnProSer 597
Db 601 CTGGAGGGAAACCATTT-ATAGAGTCTGTGAAGAACCAAGTTGGGAAACCCCAAACTCC 659
Qy 598 MetAspGlyGluLeuLeuLeuArgAlaGluGlySerThrProAlaGlyGlyLeuSer 617
Db 660 CATGAAGGGAGAACTGGTGTGAAGGACAGAGGAGCCACTTTGGCAGGCTGTGGCTTCCA 719
Qy 618 GlyGlyGlyGlyPheGlnProSerGlyLeu 627
Db 720 TG-GGTGGCGCCCTCTGGCCCTCTGTGTCTC 748

RESULT 9

AY411733

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

AY411733 2593 bp DNA linear GSS 16-DEC-2003
Pan troglodytes GPR49 gene, VIRTUAL TRANSCRIPT, partial sequence,
genomic survey sequence.
AY411733
AY411733.1 GI:39767701
GSS.
Pan troglodytes (chimpanzee)
Pan troglodytes

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pan.
1 (bases 1 to 2593)
REFERENCE
AUTHORS Clark, A.G., Glanowski, S., Nielson, R., Thomas, P., Kejarival, A.,
Todd, M.A., Tanenbaum, D.M., Civello, D.R., Lu, F., Murphy, B.,
Ferriera, S., Wang, G., Zheng, X.H., White, T.J., Sninsky, J.J.,
Adams, M.D. and Cargill, M.
TITLE
JOURNAL Science 302 (5652), 1960-1963 (2003)
PUBMED 14671302
REFERENCE
AUTHORS Clark, A.G., Glanowski, S., Nielson, R., Thomas, P., Kejarival, A.,
Todd, M.A., Tanenbaum, D.M., Civello, D.R., Lu, F., Murphy, B.,
Ferriera, S., Wang, G., Zheng, X.H., White, T.J., Sninsky, J.J.,
Adams, M.D. and Cargill, M.
TITLE
JOURNAL Direct Submission
COMMENT Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,
Rockville, MD 20850, USA
This sequence was made by sequencing genomic exons and ordering
them based on alignment.
FEATURES
source
1..2593
Location/Qualifiers
/organism="Pan troglodytes"
/mol_type="genomic DNA"
/db_xref="taxon:9598"
<1..>2593
/gene="GPR49"
/locus_tag="HCM4322"
ORIGIN
Alignment Scores:
Pred. No.: 3,916-78 Length: 2593
Score: 960.00 Matches: 209
Percent Similarity: 52.04% Conservative: 85
Best Local Similarity: 26.93% Mismatches: 191
Query Match: 29.03% Indels: 80
DB: 9 Gaps: 10
US-10-664-667-5 (1-633) x AY411733 (1-2593)
Qy 70 AspleuSerTrpAsnAlaIleArgSerIleHisProGluAlaPheSerThrLeuHisSer 89
Db 941 GATCTGCTTACAACTTATGAAGATTTA-----CCAGTTTTCAGTCTGCCAAAG 994
Qy 90 LeuValLysLeuAspLeuThrAspAenGln----- 99
Db 995 CTTGAGAAATGACCTAAGACATAATGAATC-----NNNNNNNNNNNNNNNNNN 1054
Qy 99 ----- 99
Db 1055 NNN 1114
Qy 100 -----LeuThr 101
Db 1115 NNN 1174
Qy 102 ThrLeuProLeuAlaGlyLeuGlyLeuMetHisLeuLysLeuLysGlyAsnLeuAla 121
Db 1175 TCCTTCTCTTAACCTGGGTATCATGTTAACTCACTTAAATTAACAGGAATCATGCC 1234
Qy 122 LeuSerGlnAlaPheSerLysAspSerPheProLysLeuArgIleLeuGluValProTyr 141
Db 1235 TTACAGAGCTTGATATCATCTGAAACATTTCCAGAACTCAAGGTTATAGAAATGCTTAT 1294
Qy 142 AlaTyrGlnCysCysProTyrGlyMetCysAlaSerPhePheLysAlaSerGlyGlnTrp 161
Db 1295 GCTTACCAAGTGTGTGCATTTGGAGTGTGTGAGAAATGCTATAGAAATTTCTAATCAATGG 1354
Qy 162 GluAlaGluAsp-----LeuHisLeuAspAspGluSerSerLys 175
Db 1355 AATAAAGGTGACAAACAGCAGTATGACGNNCTTCAT-----AAG 1393

Qy 176 ArgProLeuGlyLeuLeuAlaArgGlnAlaGluAenHisTyrAspGlnAspLeuAspGlu 195
Db 1394 AAAGATGCTGAATGTTTTCAGTTCAAGATGAA-----CGTGACCTTGAAGAT 1441
Qy 196 LeuGlnLeuGluMet--GluAspSerLysProHisProSerValGlnCysSerProThr 214
Db 1442 TTCCTGCTTGACTTTGAGGAAGACCTGAAAGCCCTTCATTGAGTGCATGTTCACTTCC 1501
Qy 215 ProGlyProPheLysProCysGluTyrLeuPheGluSerTrpGlyIleArgLeuAlaVal 234
Db 1502 CCAGNN---NN 1558
Qy 235 TrpAlaIleValLeuLeuSerValLeuCysAenGlyLeuValLeuLeuThrValPheAla 254
Db 1559 NNN 1618
Qy 255 GlyGlyProAlaProLeuProProValLysPheValValGlyAlaIleAlaGlyAlaAen 274
Db 1619 NNN 1678
Qy 275 ThrLeuThrGlyIleSerCysGlyLeuLeuAlaSerValAspAlaLeuThrPheGlyGln 294
Db 1679 ATGCTCACGGGAGNNNNNAGTGCCTGCTGCTGCTGGATGCTTTCACATTTTGAAGC 1738
Qy 295 PheSerGluTyrGlyAlaArgTrpGluThrGlyLeuGlyCysArgAlaThrGlyPheLeu 314
Db 1739 TTTGCACGACATGCTGCTGCTGGAGAAATGGGGTGGTTCGCATGTCAATGGTTTTTTG 1798
Qy 315 AlaValLeuGlySerGluAlaSerValLeuLeuThrLeuAlaAlaValGlnCysSer 334
Db 1799 TCCATTTTGTTCGGAATCATCTGTTTCTGCTTACTCTGGCAGCCCTCGAGCGTGGG 1858
Qy 335 ValSerValSerCysValArgAlaTyrGlyLysSerProSerLeuGlySerValArgAla 354
Db 1859 TTCTCTGTGAATATTCTGCATAAATTTGAAACAGAAAGCTCCCTTTCTAGCCTGAAGTA 1918
Qy 355 GlyValLeuGlyCysLeuAlaGlyLeuAlaAlaAlaLeuProLeuAlaSerVal 374
Db 1919 ATCATTTTGTCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1978
Qy 375 GlyGluTyrGlyAlaSerProLeuCysLeuProTyrAlaProGluGlyGlnProAla 394
Db 1979 AGCAAGTATGACCTCCCTCTCTGCTGCTGCTTGTG-----CCTTTTGGGAGCCACG 2032
Qy 395 AlaLeuGlyPheThrValAlaLeuValMetMetAsnSerPheCysPheLeuValAla 414
Db 2033 ACCATGGCTACATGGTGGCTCTCATCTGCTCAATTCCTTCTGCTTCTCATGATGACC 2092
Qy 415 GlyAlaTyrIleLysLeuTyrCysAspLeuProArgLysAspPheGluAlaValTrpAsp 434
Db 2093 ATTGCCTACACCAAGCTCTACTGCAATTTGGACAAGGGAGACCTGGAGAATATTTGGGAC 2152
Qy 435 CysAlaMetValArgHisValAlaTrpLeuIlePheAlaAspGlyLeuLeuTyrCysPro 454
Db 2153 TGCTCTATGTTAAACACATTCCTGCTGCTTCTTCCAACTGCATCCTATCTACTGCTCT 2212
Qy 455 ValAlaPheLeuSerPheAlaSerMetLeuGlyLeuPheProValThrProGluAlaVal 474
Db 2213 GTGGCTTCTTCTGCTTCTCTCTTAAATTAACCTTATCAGTCTCGTGAAGTAAT 2272
Qy 475 LysSerValLeuLeuValLeuProLeuProAlaCysLeuAsnProLeuLeuTyrLeu 494
Db 2273 AAGTTTATCTTCTGTTGGTAGTCCCACTTCTGCTGCTGCTCAATCCCTTCTCTACATC 2332
Qy 495 LeuPheAsnProHisPheArgAspLeuArgLeuArgLeuArgProArgAla----- 511
Db 2333 CTGTTCAATCTCTCACTTTAAGAGGATCTGTGAGCCCTCGGAGAAAGCAAACTTCTCTGG 2392
Qy 512 -----GlyAspSerGlyProLeuAlaTyrAlaAlaAlaGlyLeuGlyLysSer 529
Db 2393 ACAAGATCAAAACACCCAGCTTGTATGTCATTTACTCTGATGATGTCGAAAACAGTCC 2452
Qy 530 CysAspSerThrGlnAlaLeuValAlaPheSerAspValAspLeuIleLeuGluAlaSer 549

[illegible]

```
Db 980 CTGATTCGTGAAATTCACAGTGGAGCTTTGGGAAGCTTGGGACAAATTAACCTGGAT 1039
Qy 95 LeuThrAspAsnGlnLeuThrLeuProLeuAlaGlyLeuGlyGlyLeuMethHisLeu 114
Db 1040 GTGAGTTTCATGAATTAACCTCAATTCCTACGGAAGGCTGAATGGGCTCAATCAACTT 1099
Qy 115 LysLeuLysGlyAsnLeuAlaLeuSerGlnAlaPheSerLysAspSerPheProLysLeu 134
Db 1100 AAACCTGTGGGTAACTTCAGCTGAAGATGCTTGGCAGCCAGAGACTTTGCCAATCTC 1159
Qy 135 ArgIleLeuGluValProTyrAlaTyrGlnCysCysProTyrGlyMetCysAlaSerPhe 154
Db 1160 AGGTCCTCATAGTACCATAATGCTTATCAGTGTGTGTCATTTTGGGGGTGTGACTCTTAT 1219
Qy 155 PheLysAlaSerGlyGlnTrpGluAlaGluAspLeuHisLeuAspAspGluGluSer 174
Db 1220 -----GCAAAATTAAACACAGAGATACAGCCCCCAAGACACACAGTGTGACA 1267
Qy 175 LysArgProLeuGlyLeuLeuAlaArgGlnAlaGluAsnHisTyrAspGlnAspLeuAsp 194
Db 1268 AAAGAGAAGGTGCTACAGATGCAGCAAAATGCCACGAGC----- 1306
Qy 195 GluLeuGlnLeuGluMetGluAspSerLysProHisProSer-----ValGlnCysSer 212
Db 1307 -----ACTGCTGAAGTGAAGAACATAGCCAAATTAATCATCATTCATTTGTACA 1351
Qy 213 ProThrProGlyProPheLysProCysGluTyrLeuPheGluSerTrpGlyIleArgLeu 232
Db 1352 CTTCAACAGAGTGTCTTTAAGCCCTGTGATATTTACTGGGAGCTGGATGATCGCCTT 1411
Qy 233 AlaValTrpAlaIleValLeuLeuSerValLeuCysAsnGlyLeuValLeuLeuThrVal 252
Db 1412 ACAGTGTGGTTCATTTCTGTCGCTTGTCTTTTCAACCTGCTGTGTCATTTTAACAGTG 1471
Qy 253 PheAlaGlyGlyProAlaProLeuProValLysPheValValGlyAlaIleAlaGly 272
Db 1472 TTTGCG---TCTGTGTCATCATCTGCTGCTCCAAAGCTTTCATAGGCTGATTTCTGTG 1528
Qy 273 AlaAsnThrLeuThrGlyIleSerCysGlyLeuLeuAlaSerValAspAlaLeuThrPhe 292
Db 1529 TCTAACTTACTCATGGGATCTATACTGGCATCTTACTTTTCTTGATGCTGTGTCCTGG 1588
Qy 293 GlyGlnPheSerGluTyrGlyAlaArgTrpGluThrGlyLeuGlyCysArgAlaThrGly 312
Db 1589 GGCCGATTTGCTGAATTTGGCATTGTTGGTGGAAACTGGCAGCGGTGCAAGGTAGCTGG 1648
Qy 313 PheLeuAlaValLeuGlySerGluAlaSerValLeuLeuLeuThrLeuAlaAlaValGln 332
Db 1649 TCTCTGGCAGTCTTCTCTCGAGAGAGCCGATTTCTGTTAAACCTGGCAGCCGTGAA 1708
Qy 333 CysSerValSerValSerCysValArgAlaTyrGlyLysSerProSerLeuGlySerVal 352
Db 1709 AGAAGCGTTTGGAAAGATGTAATGAAATGGGAAAGAGTCACCTCGACAGTTC 1769
Qy 353 ArgAlaGlyValLeuGlyCysLeuAlaLeuAlaGlyLeuAlaAlaLeuProLeuAla 372
Db 1769 CAGGTGGCTGCCCTCGTAGCTTGTGGTGTGTCGAATAGCAGGCTGCTTCCCCCTTTT 1828
Qy 373 SerValGlyGluTyrGlyAlaSerProLeuCysLysLeuProTyrAlaProProGluGlyGln 392
Db 1829 CACGAGGGCAATATCTGTCATCACCTTGTGCTTGGCTTGGCTTT-----CTACAGGAG 1882
Qy 393 ProAlaAlaLeuGlyPheThrValAlaLeuValMetMetAsnSerPheCysPheLeuVal 412
Db 1883 ACACATCATTTAGGATTCAGTGTGACCTGTGCTTATTAACCTACTAGCATTTTTATTG 1942
Qy 413 ValAlaGlyAlaTyrIleLysLeuTyrCysAspLeuProArgGlyAspPheGluAlaVal 432
Db 1943 ATGGCCATTACTACACTAACTCTACTTCAACTTAGAGAAAGAGACCCGTCAGAAAC 2002
Qy 433 TrpAspCysAlaMetValArgHisValAlaTrpLeuIlePheAlaAspGlyLeuTyr 452
```

```
Db 2003 TCCAGCTTACATGATTAAAGCAGCTTGGCTTGGCTCATCTTCCACAAACTGCATCTTCTTC 2062
Qy 453 CysProValAlaPheLeuSerPheAlaSerMetLeuGlyLeuPheProValThrProGlu 472
Db 2063 TGCCCTGTTGATTTTCTCATTCGCACCATTCACGGCAATCTCCATCAGCCCGCAG 2122
Qy 473 AlaValLysSerValLeuLeuValValLeuProLeuProAlaCysLeuAsnProLeuLeu 492
Db 2123 ATAATGAAGTCTGTACGCTGATATTTCTCCGCTGCTGCTGCTGCTGAATCCAGTCTGT 2182
Qy 493 TyrLeuLeuPheAsnProHisPheArgAspAspLeuArgAlaGluArgProArgAlaGly 512
Db 2183 TAGCTTTTCTTCAACCCAAAGTTTAAAGACGACTGGAAGCTCCTGAAGCGGCTGTACCC 2242
Qy 512 ----- 512
Db 2243 AGGAACACGAGTCTCAGTCTCAGTCTCCATCAGCAGCAAGCGGTTGTGGGNAACAGGAT 2302
Qy 513 -----AspSerGlyProLeuAlaTyrAlaAlaAlaGlyGluLeuGluLysSerSer 529
Db 2303 TTCTACTACGACTGTGGCATGTATTCCACACTTGCAGGGTAACTGACTGTCTGTGACTGC 2362
Qy 530 CysAspSerThrGlnAlaLeuValAlaPheSerAspValAspLeuIleLeuGluAlaSer 549
Db 2363 TGTGAGTCAATTTCTTCTGACAAACAGTAGTATCGTGCAAAACACTTAATA----- 2410
Qy 550 GluAlaGlyArgProProGlyLeuGluThrTyrGlyPheProSerValThrLeuIleSer 569
Db 2411 -----AAATCGCAGACTTGTCTCTGTATTGACAGTGGCCTCT 2446
Qy 570 CysGlnGlnProGlyAla 575
Db 2447 TGCCAGAGCCAGAGGCC 2464

RESULT 11
BG916963
LOCUS 602816251P1 NCI_CGAP_Mam6 Mus musculus cDNA clone IMAGE:4935180 5',
DEFINITION mRNA sequence.
ACCESSION BG916963
VERSION BG916963.1 GI:14297439
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
REFERENCE 1 (bases 1 to 753)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: Jeffrey Green M.D.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LHAM10865 row: f column: 13
High quality sequence stop: 644.
FEATURES
Location/Qualifiers
1..753
/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="FVB/N"
/db_clones="taxon:10090"
/sex="female, virgin"
/tissue_type="infiltrating ductal carcinoma"
/dev_stage="5 months"
/lab_host="DH10B"
/clone_lib="NCI_CGAP_Mam6"
```

/note="Organ: mammary; Vector: pCMV-SPORT6; Site 1: SalI;
Site 2: NotI; Cloned unidirectionally. Primer: Oligo dt.
Library constructed by Life Technologies. Investigator
providing samples: Jeffrey Green, M.D., NIH"

ORIGIN

Alignment Scores:

Pred. No.: 2,14e-76 Length: 753
Score: 932.50 Matches: 206
Percent Similarity: 86.00% Conservatives: 9
Best Local Similarity: 82.40% Mismatches: 32
Query Match: 28.20% Indels: 7
DB: 4 Gaps: 1

US-10-664-667-5 (1-633) x BG916963 (1-753)

Qy 272 GlyAlaSerThrLeuThrGlyIleSerCysGlyLeuLeuAlaSerValAspAlaLeuThr 291
Db 10 GGTGACCCACGCGTCGGGGCATTTCTCTGTGCTCTCTGGCTCGGTGGACGCTTGACC 69
Qy 292 PheGlyGlnPheSerGluTyrGlyAlaArgTrpGluThrGlyLeuGlyCysArgAlaThr 311
Db 70 TATGTCAGTTTCGTGATGATGAGCCGCTGGAGAGCGGTCTGGGCTGCACGGCTACG 129
Qy 312 GlyPheLeuAlaValLeuGlySerGluAlaSerValLeuLeuLeuThrLeuAlaAlaVal 331
Db 130 GGCTTCTCTGGCTGTCTGGGTTTCAAGAGCGTCTGGTCTCTACACTGGCGGGCGGTG 189
Qy 332 GlnCys-SerValSerValSerCysValArgAlaTyrGlyIleYssSerProSerLeuGlySe 351
Db 190 CAGTGTCATCTCTGTGACCTGCTGCGGAGCTGACGGGAGCGCGCTGCGCTGGCAG 249
Qy 351 rValArgAlaGlyValLeuGlyCysLeuAlaLeuAla---GlyLeuAlaAlaLeuPr 370
Db 250 CGTCGGCGACGCGACCTGGGATGCTGGCTGCTGGCGGTGGCGCGCAAGCACTGCC 309
Qy 370 oLeuAlaSerValGlyGluTyrGlyValSerProLeuCysLeuProTyrAlaPro-ProG 390
Db 310 GCTGGCTCTGGTGGAGATGAGGCTCTCCCACTCTGCTGCTGCTGACGCCACACCG 369
Qy 390 luGlyGlnProAlaAlaLeuGlyPheThrValAlaLeuValMetMetAsnSerPheCysP 410
Db 370 AGGCGCGTGGCGACGCTGGCTGCTGTAGCCCTGGTGATGATGAACCTGCTGCT 429
Qy 410 heLeuValValAlaGlyAlaTyrIleYssLeuTyrCysAspLeuProArgGlyAspPheG 430
Db 430 TCTTGGTGGTGGCGGCGCTTACATCAAGCTCTACTGTGACCTGTCAAGGGGTGACTTG 489
Qy 430 luAlaValTyrAspCysAlaMetValArgHisValAlaTyrLeuIlePheAlaAspGlyL 450
Db 490 AGGCGGTGGGATGTGGCCATGTGGCCACGTGGC-TGGCTCATCTTTGCAGATGGCC 548
Qy 450 euLeuTyrCysProValAlaPheLeuSerPheAlaSerMetLeuGlyLeuPheProValT 470
Db 549 TCTCTACTGCGCGTGGCTTCTCTAGCTTTGCTCTATGCTGGCGCTCTTCTCTGTCFA 608
Qy 470 hrProGluAlaValIlySerValLeuValValLeuProLeuProAlaCysLeuAsnP 490
Db 609 C-CGCGAGGCTGTCAAGTCAGTCTCTTGGTGGTGGTGTGACTCTGCCTGACTGCTCAACC 667
Qy 490 roLeuLeuTyrLeuLeuPheAsnProHisPheArgAspLeuArgLeuArgProA 510
Db 668 CACTGCTCTACTG-CTCTTCAACCTCATTTCCGGGGATGACTT-CGGGGGCTCTGGCCAA 725
Qy 510 rgAlaGlyAspSerGlyProLeuAla 518
Db 726 GCCCTCGGAACCAAGGCCCTAGCC 751

RESULT 12

AK040883
LOCUS AK040883 3855 bp mRNA linear HTC 03-APR-2004
DEFINITION Mus musculus adult male aorta and vein cDNA, RIKEN full-length
enriched library, clone:A530037C04 product:CDNA FLJ14471 FIS, CLONE

ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM

REFERENCE
AUTHORS
TITLE
JOURNAL
MEDLINE
PUBMED

REFERENCE
AUTHORS
TITLE
JOURNAL
MEDLINE
PUBMED
REFERENCE
AUTHORS

TITLE
JOURNAL
MEDLINE
PUBMED
REFERENCE
AUTHORS

TITLE
JOURNAL
REFERENCE
AUTHORS

TITLE
JOURNAL
REFERENCE
AUTHORS

TITLE
JOURNAL

COMMENT

MAMMAL001030, WEAKLY SIMILAR TO LUTROPIN-CHORIOGONADOTROPIC HORMONE
RECEPTOR homolog [Homo sapiens], full insert sequence.
AK040883 GI:26088131
HTC; CAP trapper.
Mus musculus (house mouse)
Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1
Carninci,P. and Hayashizaki,Y.
High-efficiency full-length cDNA cloning
Meth. Enzymol. 303, 19-44 (1999)
99279253
10349636
2
Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K.,
Itoh,M., Konno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.
Normalization and subtraction of cap-trapper-selected cDNAs to
prepare full-length cDNA libraries for rapid discovery of new genes
Genome Res. 10 (10), 1617-1630 (2000)
20499374
11042159
3
Shibata,K., Itoh,M., Aizawa,K., Nagaoka,S., Sasaki,N., Carninci,P.,
Konno,H., Akiyama,J., Nishi,K., Kiteunai,T., Tashiro,H., Itoh,M.,
Sumi,N., Ishii,Y., Nakamura,S., Hazama,M., Nishine,T., Harada,K.,
Yamanoto,R., Matsumoto,H., Sakaguchi,S., Ikegami,T., Kashiwagi,K.,
Fujiwaka,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E., Wataniki,M.,
Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Matsuura,S., Kawai,J.,
Okazaki,Y., Muramatsu,M., Inoue,Y., Kita,A. and Hayashizaki,Y.
RIKEN integrated sequence analysis (RISA) system--384-format
sequencing pipeline with 384 multicapillary sequencer
Genome Res. 10 (11), 1757-1771 (2000)
20530913
11076861
4
The RIKEN Genome Exploration Research Group Phase II Team and the
FANTOM Consortium.
Functional annotation of a full-length mouse cDNA collection
Nature 409, 685-690 (2001)
5
The FANTOM Consortium and the RIKEN Genome Exploration Research
Group Phase I & II Team.
Analysis of the mouse transcriptome based on functional annotation
of 60,770 full-length cDNAs
Nature 420, 563-573 (2002)
6 (bases 1 to 3855)
Adachi,J., Aizawa,K., Akimura,T., Arakawa,T., Bono,H., Carninci,P.,
Fukuda,S., Furuno,M., Hanagaki,T., Hara,A., Hashizume,W.,
Hayashida,K., Hayatsu,N., Hiramoto,K., Hiraoka,T., Hirozane,T.,
Hori,F., Imotani,K., Ishii,Y., Itoh,M., Kagawa,I., Kasukawa,T.,
Kato,H., Kawai,J., Kojima,Y., Kondo,S., Konno,H., Kouda,M.,
Koya,S., Kurihara,C., Matsuyama,T., Miyazaki,A., Murata,M.,
Nakamura,M., Nishi,K., Nomura,K., Numazaki,R., Ohno,M., Ohsato,N.,
Okazaki,Y., Saito,R., Saitoh,H., Sakai,C., Sakai,K., Sakazume,N.,
Sano,H., Sabaki,D., Shibata,K., Shinagawa,A., Shiraki,T.,
Sogabe,Y., Tagami,M., Tagawa,A., Takahashi,F., Takaku-Akahira,S.,
Takeda,Y., Tanaka,T., Tomaru,A., Toya,T., Yasunishi,A.,
Muramatsu,M. and Hayashizaki,Y.
Direct Submission
Submitted (16-JUL-2001) Yoshihide Hayashizaki, The Institute of
Physical and Chemical Research (RIKEN), Laboratory for Genome
Exploration Research Group, RIKEN Genomic Sciences Center (GSC),
RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama,
Kanagawa 230-0045, Japan (E-mail:genome-res@ac.riken.jp,
URL:http://genome.gsc.riken.jp/, Tel:81-45-503-9222,
Fax:81-45-503-9216)
cDNA library was prepared and sequenced in Mouse Genome
Encyclopedia Project of Genome Exploration Research Group in Riken
Genomic Sciences Center and Genome Science Laboratory in RIKEN.
Division of Experimental Animal Research in Riken contributed to
prepare mouse tissues.

Please visit our web site for further details.

URL: <http://genome.gsc.riken.jp/>

URL: <http://fantom.gsc.riken.jp/>

FEATURES

source

```
1. 3855
Location/Qualifiers
  /organism="Mus musculus"
  /mol_type="mRNA"
  /strain="C57BL/6J"
  /db_xref="FANTOM DB:A530037C04"
  /db_xref="taxon:10090"
  /clone="A530037C04"
  /sex="male"
  /tissue_type="aorta and vein"
  /clone_lib="RIKEN full-length enriched mouse cDNA library"
  /dev_stage="adult"
<1. 642
misc_feature
  /note="CDNA FLJ14471 FIS, CLONE MAMMAL001030, WEAKLY
SIMILAR TO LUTROPIN-CHORIOGONADOTROPIC HORMONE RECEPTOR
homolog [Homo sapiens] (SPTR|Q96K69, evidence: FASTY,
85.84id, 92.1%length, match=637)
putative"
ORIGIN
Alignment Scores:
Pred. No.: 5,726-75 Length: 3855
Score: 929.00 Matches: 183
Percent Similarity: 88.73% Conservative: 6
Best Local Similarity: 85.92% Mismatches: 23
Query Match: 28.09% Indels: 1
DB: 3 Gaps: 0
US-10-664-667-5 (1-633) x AK040883 (1-3855)
Qy 422 CysAspLeuProArgGlyAspPheGluAlaValTrpAspCysAlaMetValArgHisVal 441
Db 3 TGTGACCTTCGGCGGCTCTGGCCAGCGGTGAGCGCGTGGGACTGGCCATCGTGGCCACGTG 62
Qy 442 AlatrPLeuIlePheAlaAspGlyLeuLeuTyrCysProValAlaPheLeuSerPheAla 461
Db 63 GCCTGGCTCATCTTTGCAGATGGCTCTCTACTGCCCGCGTGGCTTCTTCAGCTTTGCC 122
Qy 462 SerMetLeuGlyLeuPheProValThrProGluAlaValLysSerValLeuValVal 481
Db 123 TCTATGCTGGGCTCTTCCCTGTCAACCCGAGGCTGTCAAGTCAGTCTCTCTGGTGGTG 182
Qy 482 LeuProLeuProAlaCysLeuAsnProLeuLeuTyrLeuLeuPheAsnProHisPheArg 501
Db 183 CTGGCTCTGGCTGGCTGGCTCAACCCACTGCTCTACCTGCTCTTCAACCCCTCACTTCCGG 242
Qy 502 AspAspLeuArgLeuArgPro-ArgAlaGlyAspSerGlyProLeuAlaTyrAlaAl 521
Db 243 GATGACCTTCGGCGGCTCTGGCCAGCGCTCGGTCCCGAGGCCCTCTAGCTAGCTGCG 302
Qy 521 aAlaGlyGluLeuGluLysSerSerCysAspSerThrGlnAlaLeuValAlaPheSerAs 541
Db 303 AGCCGCTGAGCTGGAGAGAGCTCTGGCACTCCACCAAGCGCTGGTGGCTTCTCAGA 362
Qy 541 pValAspLeuLeuLeuGluAlaSerGluAlaGlyArgProProGlyLeuGluThrTyrGl 561
Db 363 TGTGATCTTATCTTGGAAAGCTTCTGAGGCTGGGCGAGCTCTCTGGCTAGAGACCTATGG 422
Qy 561 yPheProSerValThrLeuIleSerCysGlnGlnProGlyAlaProArgLeuGluGlySe 581
Db 423 CTTCCTTCAGTGACCTCATCTCCGACATCAGCGGGGGCTACACGCTGGAGGGAAA 482
Qy 581 rHisCysValGluProGluGlyAsnHisPheGlyAsnProGlnProSerMetAspGlyGl 601
Db 483 CCATTTTGTAGAGTCTGATGGAAACCAAGTTTGGGAACCCACAACCTCCCATGAGGGAGA 542
Qy 601 uLeuLeuLeuArgAlaGluGlySerThrProAlaGlyGlyGlyLeuSerGlyGlyGl 621
Db 543 ACTGCTGTGAAGGAGGGAGGCCACTTTGGCAGGCTGTGGCTCTCTCCGTTGGTGGAGC 602
```

Qy 621 yPheGlnProSerGlyLeuAlaPheAlaSerHisVal 633

Db 603 CCTCTGGCCCTCTGGCTCTCTCTTTTGCCTCTCACTTG 639

RESULT 13

BB636845

LOCUS

DEFINITION

BB636845

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

BB636845 690 bp mRNA linear EST 26-OCT-2001
BB636845 RIKEN full-length enriched, adult male aorta and vein Mus
musculus cDNA clone A530037C04 5', mRNA sequence.

BB636845

BB636845.1 GI:16472635

EST.

Mus musculus (house mouse)

Mus musculus

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 690)

Arakawa, T., Carninci, P., Fukuda, S., Furuno, M., Hanagaki, T.,

Hara, A., Hiramoto, K., Hori, F., Ishii, Y., Ito, M., Kawai, J.,

Konno, H., Kouda, M., Koya, S., Matsuyama, T., Miyazaki, A., Nomura, K.,

Ohno, M., Okazaki, Y., Okido, T., Saito, R., Sakai, C., Sakai, K.,

Sano, H., Sasaki, D., Shibata, K., Shinagawa, A., Shiraki, T.,

Sogabe, Y., Suzuki, H., Tagami, M., Tagawa, A., Takahashi, F.,

Takeda, Y., Tanaka, T., Toya, T., Muramatsu, M. and Hayashizaki, Y.

RIKEN Mouse ESTs (Arakawa, T., et al. 2001)

Unpublished (2001)

Contact: Yoshihide Hayashizaki

Laboratory for Genome Exploration Research Group, RIKEN Genomic

Sciences Center (GSC), Yokohama Institute

The Institute of Physical and Chemical Research (RIKEN)

1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan

Tel: 81-45-503-9222

Fax: 81-45-503-9216

Email: genome-rsgsc.riken.jp, URL: <http://genome.gsc.riken.jp/>

Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,

Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.

Normalization and subtraction of cap-trapper-selected cDNAs to

prepare full-length cDNA libraries for rapid discovery of new

genes. Genome Res. 10 (10), 1617-1630 (2000)

wagi, K., Fujiwaki, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E.,

Matsuura, S., Kawai, J., Ishikawa, T., Ozawa, K., Tanaka, T.,

and Hayashizaki, Y.

RIKEN integrated sequence analysis (RISA) system--384-format

sequencing pipeline with 384 multicapillary sequencer. Genome Res.

10 (11), 1757-1771 (2000)

Konno, H., Fukunishi, Y., Shibata, K., Itoh, M., Carninci, P.,

Sugahara, Y. and Hayashizaki, Y.

Computer-based methods for the mouse full-length cDNA

encyclopedia: real-time sequence clustering for construction of a

nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)

Kondo, S., Shinagawa, A., Saito, T., Kiyosawa, H., Yamanaka, I.,

Aizawa, K., Fukuda, S., Hara, A., Itoh, M., Kawai, J., Shibata, K. and

Hayashizaki, Y.

Computational Analysis of Full-Length Mouse cDNAs Compared with

Human Genome Sequences. Mamm. Genome. 12, 673-677 (2001)

Please visit our web site (<http://genome.gsc.riken.go.jp>) for

further details.

e mouse tissues.

FEATURES

source

Location/Qualifiers

1. 690

/organism="Mus musculus"

/mol_type="mRNA"

/db_xref="taxon:10090"

/clone="A530037C04"

/sex="male"

/tissue_type="aorta and vein"

/dev_stage="adult"

/lab_host="DH10B"

/clone_lib="RIKEN full-length enriched, adult male aorta

and vein"

/note="Site 1: SalI; Site 2: BamHI; cDNA library was

prepared and sequenced in Mouse Genome Encyclopedia

Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN, Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. 1st strand cDNA was verified with a primer [5', GAGAGAGAGAGATCCAGAGACTCTTTTTTTTTTTTNN 3'], cDNA was prepared by using trehalose thermo-activated reverse transcriptase and subsequently enriched for full-length by cap-trapper. cDNA went through one round of normalization to Rot = 20.0 and subtraction to Rot = 459.0. Second strand cDNA was prepared with the primer adapter of sequence [5', GAGAGAGATTCGAGTAAATTAATCCCCCCOCCCC 3']. cDNA was cleaved with XhoI and BamHI. Vector: a modified pBluescript KS(+) after bulk excision from Lambda FLX 1."

ORIGIN

Alignment Scores:
Pred. No.: 1e-74 Length: 690
Score: 914.00 Matches: 181
Percent Similarity: 88.68% Conservative: 7
Best Local Similarity: 85.38% Mismatches: 24
Query Match: 27.64% Indels: 1
DB: 2 Gaps: 0

US-10-664-667-5 (1-633) x BB636845 (1-690)

Qy 422 CysAspLeuProArgGlyAspPheGluAlaValTrpAspCysAlaMetValArgHisVal 441
Db 1 TGTGACCTGCCACGGGTGACTTTGAGCGCGTGTGGAGCTGCGCATGGTGGCGCAGGTG 60
Qy 442 AlaTrpLeuIlePheAlaAspGlyLeuLeuTyrCysProValAlaPheLeuSerPheAla 461
Db 61 GCCTGGCTCATCTTGCAGATGGCTCTCTACTGCCCGCGGTGCTTCTCAGCTTTGCC 120
Qy 462 SerMetLeuGlyLeuPheProValThrProGluAlaValLysSerValLeuValVal 481
Db 121 TCTATGTGGGCTCTTCCCTGTACCCCGGAGGCTGTCAAGTCACTCTCTGGTGGTG 180
Qy 482 LeuProLeuProAlaCysLeuAsnProLeuLeuTyrLeuLeuPheAsnProHisPheArg 501
Db 181 CTGGCTCTGGCTGGCTGCTCAAGCCACTGCTTACCTGCTCTTCAACCCCTCACTCCGG 240
Qy 502 AspAspLeuArgArgLeuArgProArgAlaGlyAspSerGlyProLeuAlaTyrAlaAla 521
Db 241 GATGACCTTCGGCGGCTCTGGCCAGCCCTCGTCCCGGAGGCGCCCTAGCCCTAGCGTGA 300
Qy 522 AlaGlyGluLeuGluLysSerSerCysAspSerThrGlnAlaLeuValAlaPheSerAsp 541
Db 301 GCCGTGAGCTGGAGAGAGCTCTCGGACTCCACCAAGCGCTGGTGGCTTTCTCAGAT 360
Qy 542 ValAspLeuIleLeuGluAlaSerGluAlaGlyArgProGlyLeuGluThrTyrGly 561
Db 361 GTGGATCTTATCTGGAAGCTTCTGAGGCTGGGAGCGCTCTCGGCTAGAGACCTATGGC 420
Qy 562 PheProSerValThrLeuIleSerCysGlnGlnProGlyAlaProArgLeuGluGlySer 581
Db 421 TTCCCTTCAGTACCCCTCATCTCCGACATCAGCCGGGGGTACACAGGCTGGAGGGAAC 480
Qy 582 HisCysValGluProGluGlyAsnHisPheGlyAsnProGlnProSerMetAspGlyGlu 601
Db 481 CATTTTGTAGAGTCTGATGGAAACCAAGTTTGGGAACCCACCACTCCCATGAAGGGAGAA 540
Qy 602 LeuLeuLeuArgAlaGluGlySerThrProAlaGlyGlyGlyLeuSerGlyGlyGly 621
Db 541 CTGCTGCTGAAGGAGGAGGAGCCACTTTGGCAGGCTGTGGCTCTTCCGTGGGTGGAGCC 600
Qy 622 PheGlnProSerGlyLeuAlaPheAlaSerHisVal 633
Db 601 CTCTGGCCCTCTGGCTC-TCTCTTGGCTCTCTCATTG 635

RESULT 14
BF158974

LOCUS

DEFINITION 601766476P1 NCI_CGAP_Lu29 Mus musculus cDNA clone IMAGE:3982506 5', mRNA sequence.

ACCESSION BF158974

VERSION BF158974.1 GI:11039068

KEYWORDS EST.

SOURCE Mus musculus (house mouse)

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. 1 (bases 1 to 869)

AUTHORS NIH-MGC http://mgs.nci.nih.gov/.

TITLE National Institutes of Health, Mammalian Gene Collection (MGC)

JOURNAL Unpublished (1999)

COMMENT Contact: Robert Strausberg, Ph.D.

Email: cgabs-remail.nih.gov

Tissue Procurement: Gilbert Smith, Ph.D.

cDNA Library Preparation: Life Technologies, Inc.

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Incyte Genomics, Inc.

Clone distribution: MGC clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at:

http://image.llnl.gov

Plate: L1AM9181 row: g column: 19

High quality sequence stop: 640.

FEATURES

Location/Qualifiers

1..869

/organism="Mus musculus"

/mol_type="mRNA"

/strain="CZECH II"

/db_xref="taxon:10090"

/clone="IMAGE:3982506"

/tissue_type="spontaneous tumor, metastatic to mammary."

/lab_host="DH10B"

/clone_lib="NCI CGAP Lu29"

/notes="Organ: lung; Vector: pCMV-SPORT6; Site 1: SalI;

Site 2: NotI; Cloned unidirectionally. Primer: Oligo dT.

Library constructed by Life Technologies. Investigator

providing samples: Gilbert Smith, NIH"

ORIGIN

Alignment Scores:

Pred. No.: 4.3e-72 Length: 869
Score: 887.50 Matches: 212
Percent Similarity: 77.78% Conservative: 19
Best Local Similarity: 71.38% Mismatches: 39
Query Match: 26.84% Indels: 32
DB: 2 Gaps: 4

US-10-664-667-5 (1-633) x BF158974 (1-869)

Qy 175 LysArgProLeuGlyLeuLeuAlaArgGlnAlaGluAsnHisTyrAspGlnAspLeuAsp 194
Db 2 AAGAGGCCCTTGGGTCTCTTCTTGGCAAGCTGAGAACCACTATGACTA-GACCTGGAT 60
Qy 195 GluLeuGlnLeuGluMetGluAspSerLysProHisProSerValGlnCysSerProThr 214
Db 61 GAGCTCCAGATGGGGAGAGGAGCTCAAGCCAAACCCAGTGTCCAGTGCAGCCCTGTT 120
Qy 215 ProGlyProPheLysProCysGluTyrLeuPheGluSerTrpGlyIleArgLeuAlaVal 234
Db 121 CCAGGCCCTTCAAGCCCTCGAGACCTCTTTTGGAGAGCTGGGGCATCCGCTTGTGTG 180
Qy 235 TrpAlaIleValLeuLeuSerValLeuCysAsnGlyLeuValLeuLeuThrValPheAla 254
Db 181 TGGGCCATCTGCTGCTCTCCGTACTCTGTAAAGGGCTGGTGTGCTGCTGACAGTCTTTGCC 240
Qy 255 GlyGlyProAlaProLeuProValLysPheValValGlyAlaIleAlaGlyAlaAsn 274
Db 241 AGCGGACCCAGCCCGTGTCCCGCTCAAGCTTGTGGTGGTGGATGGCAGGCGGCAAC 300
Qy 275 ThrLeuThrGlyIleSerCysGlyLeuLeuAlaSerValAlaLeuThrPheGlyGln 294

301	Db		GCCTGACGGGATTTCCTGTGGTCTCTCCGCTCGGTGGACGCGCTTGACCTATGTGTCAG	360
295	Qy		PheSerGluTyrGlyAlaArgTrpGluThrGlyLeuGlyCysArgAlaThrGlyPheLeu	314
361	Db		TTCCGCTGAGATGGAGCCCGCTGGGAGAGCGGTCTGGGCTGCCAGGCTACGGGCTTCCTTG	420
315	Qy		AlaValLeuGlySerGluAlaSerValLeuLeuLeuThrLeuAlaValGlnCysSer	334
421	Db		GCTGTCTGGGTTACAGAGGGTCGGTGTGTGCTCACATCTGGGGGGGTGCGATGTCAGC	480
335	Qy		ValSerValSerCysValArgAlaTyrGlyIySerProSerLeuGlySerValArgAla	354
481	Db		ATCTCTGTGACCTG-GTCCGACCTACGGGAAGCGCGCTGCT-GGCACGCTCCGCGCA	538
355	Qy		GlyValLeuGlyCysLeuAlaLeuAlaGlyLeuAlaAlaLeuProLeuAlaSerVal	374
539	Db		GGCGCACTCTGATGCTCGCGCTGGCGGGCTGGCGCTGCAGCACTGCG-CTGGGCTCAGTG	597
375	Qy		GlyGluTyrGlyAlaSerProLeuCysLeuProTyrAlaPro-ProGluGlyGlnProAl	394
598	Db		GGAGAGTATGGCGCTCCCCATCTGACGTGCTAAGCCCCACCCCGAGTGGCGGGCGGC	657
394	Qy		alaLeu-GlyPheThrValAlaLeuValMetMet-----AsnSerPheCysPheL	411
658	Db		GGCCCTGGGGCTTCGCTGTAGACCTTGTGATTGATTGACATCCGCTCTGGCTCCCTGG	717
411	Qy		euValValAlaGlyAlaTyrIleLys-LeuTyrCysAspLeuProArgGlyAspPheGlu	430
718	Db		TTGGTGGGCGGGCGCCTCAATCAACGCTCTACTGTGACCTGCCCGG-GGTGACTTGAGG	776
431	Qy		AlaVal-----TrpAspCysAlaMetVal	438
777	Db		CCGTTGGGATGCCATGTGGCCGGTGGCTGGGCCATCTTTGGGAT-----	824
439	Qy		ArgHisValAlaTrpIlePheAlaAspGlyLeuLeuTyrCys	453
825	Db		-----TGGCTCTTATGGC-----GGCCTCTCGCATTCG	854

RESULT 15	CN258480	680 bp	linear	EST 16-MAY-2004
LOCUS	CN258480			
DEFINITION	17000532179685 GRN_ES Homo sapiens	cdna 5',	mRNA sequence.	
ACCESSION	CN258480			
VERSION	CN258480.1	GI:47274894		
KEYWORDS	EST.			
SOURCE	Homo sapiens (human)			

ORGANISM Homo sapiens (human)

REFERENCE Zukayoka; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo. 1 (bases 1 to 680)

AUTHORS Brandenberger, R., Wei, H., Zhang, S., Lei, S., Murage, J., Fisk, G.J., Li, Y., Xu, C., Fang, R., Gugler, K., Rao, M.S., Mandalam, R., Lebkowski, J and Stanton, L.W.

TITLE Transcriptome characterization elucidates signaling networks that control human ES cell growth and differentiation

JOURNAL Nat. Biotechnol. 22 (6), 707-716 (2004)

COMMENT Contact: Brandenberger R

Geron Corporation
230 Constitution Drive, Menlo Park, CA 94025, USA
Tel: 650 473 8658
Fax: 650 473 7760
Email: rbrandenberger@geron.com
Insert Length: 680 Std Error: 0.00.

```

FEATURES
source
1.680
Location/Qualifiers
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/tissue_type="embryonic stem cells, cell lines H1, H7, and H9"
/clone_lib="GRN ES"

```

```

/note="oligo dt primed, full-length enriched cDNA library
from undifferentiated hES cell lines H1 (p32), H7 (p29),
and H9 (p26) maintained in feeder-free conditions"

ORIGIN

Alignment Scores:
Pred. No.:      2.33e-64      Length:      680
Score:          803.00        Matches:     171
Percent Similarity: 89.11%    Conservative: 9
Best Local Similarity: 84.65%  Mismatches:  16
Query Match:     24.28%      Indels:      8
DB:              7           Gaps:        2

```

US-10-664-667-5 (1-633) x CN258480 (1-680)

Qy	12	CysArgTyrArg-----SerGlyIleProGlySer-ThrHisAlaSerValGluAr	28
Db	79	TGCTCCCATGGGGATGTGCCAACAGCTGCCAGGCTCCGAGTCTCGGAACATGTCTCA	138
Qy	28	gSerGln--GlyLeuSerLeuProAlaHisProAlaSerLeuAlaLeuAlaAlaSe	47
Db	139	ATCAAAATTGAGGAGCTCCCGAGCTGCACAGGTGCAGAAATTGGAGGAAA-TCGCGCTC	197
Qy	47	rAsnThrThrAlaSerGlyLysLeuGlu**AspThrPheSerGlnLeuSerSerLeuGl	67
Db	198	CAACACAACCGCATCTGGGAAATTGGAGCT-GACACCTTCAGCCAGCTGAGCTCCCTGCA	256
Qy	67	nAlaLeuAspLeuSerTrpAsnAlaIleArgSerIleHisProGluAlaPheSerThrIle	87
Db	257	AGCCCTTGGATCTTAGCTGGAAAGCCATCCGGTCCATCCACCCGAGGCGCTTCTCCACCCCT	316
Qy	87	uHiserLeuValLysLeuAspLeuThrAspAsnGlnLeuThrThrLeuProLeuAlaGl	107
Db	317	GCATCTCCCTGGTCAAGCTGGACCTGCACGACACCAACGAGCTGACCACACTGCCCTGGCTGG	376
Qy	107	yLeuGlyGlyLeuMethIstLeuLysLeuLyGlyAsnLeuAlaLeuSerGlnAlaPheSe	127
Db	377	ACTTGGGGGCTTGATGCATCTGAAGCTCAAAGGGAACCTTGCTCTCTCCAGGCGCTTCTC	436
Qy	127	rLysAspSerPheProLysLeuArgIle-LeuGluValProTyrAlaTyrGlnCysCvp	147
Db	437	CAAGGACAGTTTCCCAAACATGAGGATCCTTGGAGTGCCCTTATGCCATCCAGTGTCTGTC	496
Qy	147	roTyrGlyMetCysAlaSerPhePheLysAlaSerGlyGlnTrpGluAlaGluAspLeuH	167
Db	497	CCTATGGATGTGTGCCAGCTTCTTCAAGGCTCTCTGGGCAGTGGGAGGCTGAAGACCTTC	556
Qy	167	isLeuAspAspGluGluSerSerLysArgProLeuGlyLeuLeuAlaArgGlnAlaGluA	187
Db	557	ACCTTGTATGATGAGGAGTCTTCAAAAAGGCCCTTGGGCTCTTGTGCCACAGACAGAGA	616
Qy	187	snHisTyrAspGlnAspLeuAspGluLeuGlnLeuGluMetGluAspSerLysProHisP	207
Db	617	ACCACTATGACCAGGACCTGGATGAGCTCCAGCTGGAGATGGAGGACTCAAGGCCACACC	676
Qy	207	ro	207
Db	677	CC	678

Search completed: April 14, 2005, 21:48:37
Job time : 3652 secs

This Page Blank (uspto)

i LOCATION: (1)...(2484)
US-09-799-451-723

Alignment Scores:

Pred. No.: 6,59e-280 Length: 2988
Score: 3040.00 Matches: 594
Percent Similarity: 97.41% Conservative: 7
Best Local Similarity: 96.27% Mismatches: 14
Query Match: 91.93% Indels: 4
DB: 4 Gaps: 1

US-10-664-667-5 (1-633) x US-09-799-451-723 (1-2988)

QY 19 ProGlySer-ThrHisAlaSerValGluArgSerGln---GlyLeuSerLeuProAlaHi 37
DB 638 CCAGGCTCCGAGTCTCGAAGTGTCTCAAAATTTGAGGAGCTGCCAGCTGCACA 697
QY 37 sProAlaSerLeuAlaLeuAlaAlaSerAsnThrThrAlaSerGlyLysLeuGlu** 57
DB 698 GGTGTGAGAAATTTGAGGAGAA-TGGGCTCCAAACACCCGATCTGGAAATTTGGAGCT 756
QY 57 *AspThrPheSerGlnLeuSerSerLeuGlnAlaLeuAspLeuSerTrpAsnAlaIleAr 77
DB 757 -GACACCTTCAGCAGCTGAGCTCCCTGCAAGCCCTGGATCTTAGCTGGAACGCCATCCG 815
QY 77 gSerIleHisProGluAlaPheSerThrLeuHisSerLeuValLysLeuAspLeuThrAs 97
DB 816 GTCCATCCACCTGAGGCTTTCTCACCTTCGATCTCCCTGCTGAGCTGGACCTGCACA 875
QY 97 pAsnGlnLeuThrThrLeuProLeuAlaGlyLeuGlyLeuMethHisLeuLysLeuLy 117
DB 876 CAACGAGCTGACACATGCTCCCTTGGCTGGACTTTGGGGCTTGTATGATCTGAAGCTCAA 935
QY 117 sGlyAsnLeuAlaLeuSerGlnAlaPheSerLysAspSerPheProLysLeuArgIleLe 137
DB 936 AGGGAACCTTCTCTCCAGGCTTCTCCAAAGCACAGTTTCCCAAACTGAGGATCCT 995
QY 137 uGluValProTyraLysGlnCysCysProTyrglyMetCysAlaSerPheLysAl 157
DB 996 GGAGGTGCTTATGCTTACCATGCTGCTGCTATGGAGTGTGCTGAGCTTCTTCAAGGC 1055
QY 157 aserGlyGlnTrpGluAlaGluAspLeuHisLeuAspGluGluSerLysArgPr 177
DB 1056 CTCTGGGAGTGGAGGCTGAGACCTTCACCTTGTATGATGAGGAGCTTCCAAAAAGGCC 1115
QY 177 oLeuGlyLeuLeuAlaArgGlnAlaGluAsnHisTyraSpGlnAspLeuAspGluLeuGl 197
DB 1116 CTTGGGCTCTTCCAGACAAGCAGAGAACCACTATGACACGAGACCTGGATGAGCTCCA 1175
QY 197 nLeuGluMetGluAspSerLysProHisProSerValGlnCysSerProThrProGlyPr 217
DB 1176 GCTGGAGATGAGGAGCTCAAGAGCCACACCCAGTGTCCAGTGTAGCCCTACTCCAGGCC 1235
QY 217 oPheLysProCysGluTyraLeuPheGluSerTrpGlyLeuArgLeuAlaValTtpAlaIl 237
DB 1236 CTTCAAGCCCTGTAGTACCTCTTTTGAAGCTGGGCAATCCGCTTGGCCGTGGGCCAT 1295
QY 237 eValLeuLeuSerValLeuCysAsnGlyLeuValLeuLeuThrValPheAlaGlyGlyPr 257
DB 1296 CGTGTGCTCTCCGTCCTGCAATGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1355
QY 257 oAlaProLeuProProValLysPheValValGlyAlaIleAlaLeuAlaLeuThrLeuTh 277
DB 1356 TGTCCCCCTGCCCGGTCAAGTTTGTGGTAGGTGCGATTCAGCGGCCAACACCTTGAC 1415
QY 277 rGlyLysSerCysGlyLeuLeuAlaSerValAspAlaLeuThrPheGlyGlnPheSerGl 297
DB 1416 TGGATTTCTGTGGCTTCTAGCTCAGTCGATGCCCTGACCTTTGGTCAGTTCTCTGA 1475
QY 297 uTyrglyAlaArgTrpGluThrGlyLeuGlyCysArgAlaThrGlyPheLeuAlaValLe 317
DB 1476 GTACGGAGCCGCTGGAGAGCGGGCTAGGTGCTGCGGGCCACTGGCTTCTTGGCAGTACT 1535

QY 317 uGlySerGluAlaSerValLeuLeuLeuThrLeuAlaAlaValGlnCysSerValSerVa 337
DB 1536 TGGGTCCGAGGATCGGTGTCTGTCTCACTCTGCGCGCAGTGAGTGCGAGCGTCTCGGT 1595
QY 337 lSerCysValArgAlaTyrglyLysSerProSerLeuGlySerValArgAlaGlyValLe 357
DB 1596 CTCCTGTGTCCGGGCTTATGGAGAGTCCCTCCCTCCCTGGCAGCGTTCAGCAGGGGTCT 1655
QY 357 uGlyCysLeuAlaLeuAlaGlyLeuAlaAlaLeuProLeuLeuAspLeuGlyGlyTy 377
DB 1656 AGGCTGCTGCGCATGCGAGGCTGGCGCGCGCATCTCCCTGCGCTCAGTGGGAGATA 1715
QY 377 rGlyAlaSerProLeuCysLeuProTyraAlaProProGluGlyGlnProAlaAlaLeuGl 397
DB 1716 CGGGGCTTCCCACTTCTGCTTCCCTACCGCCACCTGAGGCTGAGCAGCGAGCCCTGGG 1775
QY 397 yPheThrValAlaLeuValMetMetAsnSerPheCysPheLeuValValAlaGlyAlaTy 417
DB 1776 CTTACCGCTGGCCCTGGTGATGATGAATCTCTTCTGTTTCTGCTGCGCGGTGCTA 1835
QY 417 rIleLysLeuTyrcysAspLeuProArgGlyAspPheGluAlaValTrpAspCysAlaMe 437
DB 1836 CATCAAACTGTACTGTGACCTGCGCGGGCGACTTTGAGGCGGTGGGACTCGGCAT 1895
QY 437 tValArgHisValAlaTrpLeuLeuPheAlaAspGlyLeuLeuTyrcysProValAlaPh 457
DB 1896 GGTGAGCAGCTGGCTTGGCTCATCTTCGAGAGCGGGCTCTCTACTGTCCCGTGGCTT 1955
QY 457 eLeuSerPheAlaSerMetLeuGlyLeuPheProValThrProGluAlaValLysSerVa 477
DB 1956 CTTACGCTTCCCTCCATGCTGGGCTTCTTCCCTGTTCACGCCGAGCGCTCAAGTCTGT 2015
QY 477 lLeuLeuValValLeuProLeuProAlaCysLeuAsnProLeuLeuTyrcysLeuPheAs 497
DB 2016 CTTGCTGTGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2075
QY 497 nProHisPheArgAspLeuArgLeuArgLeuArgLeuArgLeuArgLeuArgLeuArgLe 517
DB 2076 CCCCCCTTCGGGATGACCTTTCGGGCTTTCGGGCTTTCGGGCTTTCGGGCTTTCGGG 2135
QY 517 uAlaTyraAlaAlaGlyGluLeuGluLysSerSerCysAspSerThrGlnAlaLeuVa 537
DB 2136 AGCTATGCTGCGCGGGGAGCTGGAGAGAGCTCTGTGTATCTTACCAGGCGCTTGT 2195
QY 537 lAlaPheSerAspValAspLeuLeuLeuAlaSerGluAlaGlyArgProGlyLe 557
DB 2196 AGCCTTCTCTGATGTGATCTCATTTCTGGAAGCTTCTGAAGCTGGGCGGCGCTTGG 2255
QY 557 uGluThrTyrglyPheProSerValThrLeuLeuLeuSerCysGlnGlnProGlyAlaProAr 577
DB 2256 GAGACCTATGGCTTCCCTTCAGTGACCTCATCTCTGTGAGCAGCGGCGGCCCCAG 2315
QY 577 gLeuGluGlySerHisCysValGluProGluGlyAsnHisPheGlyAsnProGlnProSe 597
DB 2316 GCTGGAGGCGAGCATTTGTAGAGCAGAGGGGAGACCTTTGGGAGACCCCAACCTC 2375
QY 597 rMetAspGlyGluLeuLeuLeuArgAlaGluGlySerThrProAlaGlyGlyLeuSe 617
DB 2376 CATGGATGGAGAACTGTGCTGAGGCGAGAGGATCTACGCCAGCAGGTGGAGCTTGTG 2435
QY 617 rGlyGlyGlyGlyPheGlnProSerGlyLeuAlaPheAlaSerHisVal 633
DB 2436 AGGGGTGGGCGCTTTCAGCCCTTCTGCGCTTGGCTTGTGCTTACACGTG 2484

RESULT 2

US-09-170-496D-263
; Sequence 263, Application US/09170496D
; Patent No. 655339
; GENERAL INFORMATION:
; APPLICANT: Behan, Dominic P.
; APPLICANT: Chalmers, Derek T.
; APPLICANT: Liaw, Chen W.
; TITLE OF INVENTION: No. 6555339-Endogenous, Constitutively Activated Human G Protein-

;; TITLE OF INVENTION: Receptors
;; FILE REFERENCE: AREN-0040
;; CURRENT APPLICATION NUMBER: US/09/170,496D
;; CURRENT FILING DATE: 1998-10-13
;; NUMBER OF SEQ ID NOS: 294
;; SOFTWARE: PatentIn version 3.1
;; SEQ ID NO 263
;; LENGTH: 2724
;; TYPE: DNA
;; ORGANISM: Homo sapiens
US-09-170-496D-263

Alignment Scores:

Pred. No.:	1,99e-111	Length:	2724
Score:	1275.00	Matches:	250
Percent Similarity:	64.92%	Conservative:	96
Best Local Similarity:	46.90%	Mismatches:	157
Query Match:	38.55%	Indels:	30
DB:	4	Gaps:	8

US-10-664-667-5 (1-633) x US-09-170-496D-263 (1-2724)

Qy	54	LysLeuGlu***AspThrPheSerGlnLeuSerSerLeuGlnAlaLeuAspLeuSerTrp	73
Db	1162	GAATTAAGTTGACACTTCCAGCAGTTGCTTAGCCTCCGATCGCTGAATTTGGCTTGG	1221
Qy	74	AsnAlaIleArgSerIleHisProGluAlaPheSerThrLeuHisSerLeuValIysLeu	93
Db	1222	AACAAATGCTATTATTACCCCAATGCAATTTCCACTTTGCCATCCCTAATAAGCTG	1281
Qy	94	AspLeuThrAspAsnGlnLeuThrThrLeuProLeuAlaGlyLeuGlyGlyLeuMethis	113
Db	1282	GACCTATCGTCCAACTCCTGCTGCTTTCTATTAACCTGGGTTACATGTTTAACAC	1341
Qy	114	LeuIysLeuIysGlyAsnLeuAlaLeuSerGlnAlaPheSerLysAspSerPheProlys	133
Db	1342	TTAAATTAACAGGAATCATGCCCTTACAGAGTTGATCATCTGAAAACCTTTCCAGAA	1401
Qy	134	LeuArgIleLeuGluValProTyrAlaTyrGlnCysCysProTyrGlyMetCysAlaSer	153
Db	1402	CTCAAGGTTATAGAAATGCTTTATGCTTACAGTGCTGTGATTTGGAGTGTGAGNAT	1461
Qy	154	PhePheIysAlaSerGlyGlnTrp-----GluAlaGluAspLeuHis	167
Db	1462	GCCTATAAGATTTCTAATCAATGGAATAAAGGTGACAAACAGCAGTATGGACGACCTTCAT	1521
Qy	168	LeuAspAspGluGluSerSerLysArgProLeuGlyLeuLeuAlaAArgGlnAlaGluAsn	187
Db	1522	-----AAGAAAGATGCTGGAATGTTTCAGGCTCAAGATGAA---	1557
Qy	188	HisTyrAspGlnAspLeuGluLeuGlnLeuMet---GluAspSerLysProHis	206
Db	1558	-----CGTGACCTTGAGATTTCTGCTTGATTTTGAGGAAGACCTGAAAGCCCTT	1608
Qy	207	ProSerValGlnCysSerProGlyProGlyProPheIysProCysGlyTyrLeuPheGlu	226
Db	1609	CATTGAGTGCAGTGTTCACCTTCCAGGCCCTTCAAAACCTGTGGAACACCTGCTTGAT	1668
Qy	227	SerTrpGlyIleArgLeuAlaValTrpAlaIleValLeuLeuSerValLeuCysAsnGly	246
Db	1669	GGCTGGCTGATCAGAAATGGAGTGGACCATAGCAGTCTGGCAGCTTACTTGAATGCT	1728
Qy	247	LeuValLeuLeuThrValPheAlaGlyGlyProAlaProLeuProValIysPheVal	266
Db	1729	TTGGTGACTTCAACAGTTTTCAGA---TCCCTCTGTATCATTTCCCATTAACCTGTTA	1785
Qy	267	ValGlyAlaIleAlaGlyAlaAsnThrLeuThrGlyIleSerCysGlyLeuLeuAlaSer	286
Db	1786	ATTGGGGTCAATCGCAGCATGAATGCTACGGGAGTCTCCAGTGGCGTCTGGCTGGT	1845
Qy	287	ValAspAlaLeuThrPheGlyGlnPheSerGluTyrGlyAlaAArgTrpGluThrGlyLeu	306
Db	1846	GTGGATGCTTCACTTTTGGCAGCTTTTGCACGACATGTTGCTGTTGGAGAAATGGGGT	1905

Qy	307	GlyCysArgAlaThrGlyPheLeuAlaValLeuGlySerGluAlaSerValLeuLeuLeu	326
Db	1906	GGTTGCCATGTCATTGGTTTTTTCATTTTTCAGAAATCATCTGTTTCTGCTT	1965
Qy	327	ThrLeuAlaAlaValGlnCysSerValSerValSerCysValArgAlaTyrGlyIysSer	346
Db	1966	ACTCTGGCAGCCCTGGAGCGTGGTCTCTCTGGAATATTCTGCAAAATTTGAAACGAAA	2025
Qy	347	ProSerLeuGlySerValArgAlaGlyValLeuGlyCysLeuAlaLeuAlaGlyLeuAla	366
Db	2026	GCTCCATTTTCTAGCCTGAAAGTAATCATTTTGTCTGTGCCCTGTGGCTTGTACCATG	2085
Qy	367	AlaAlaLeuProLeuAlaSerValGlyGluTyrGlyAlaSerProLeuCysLeuProTyr	386
Db	2086	GCCGAGTTCCCTGCTGGGTGGCAGCAAGTATGGCGCTCCCTCTCTGCTGCTGCTTGT	2145
Qy	387	AlaProProGluGlyGlnProAlaLeuAlaLeuGlyPheThrValAlaLeuValMetMetAsn	406
Db	2146	-----CCTTTTGGGAGCCCGACCATGGCTGACATGGTGGCTCTCATCTTGTCTCAAT	2199
Qy	407	SerPheCysPheLeuValValAlaGlyAlaTyrIleLysLeuTyrCysAspLeuProArg	426
Db	2200	TCCCTTTGCTTCTCATGATGACCATTTGCGCTACACCAAGCTCTACTGCAATTTGGACAAG	2259
Qy	427	GlyAspPheGluAlaValAlaTrpAspCysAlaMetValArgHisValAlaTrpLeuIlePhe	446
Db	2260	GGAGACCTGGAGATAATTTGGGACTGCTCTATGTTGTAACACACATTTGCCCTGTGCTCTC	2319
Qy	447	AlaAspGlyLeuLeuTyrCysProValAlaPheLeuSerPheAlaSerMetLeuGlyLeu	466
Db	2320	ACCAACTGCATCTCTAAACTGCGCTGTGGCTTTCTTGTCTCTCTCTCTCTCTCTCTCT	2379
Qy	467	PheProValThrProGluAlaValLysSerValLeuLeuValValLeuProLeuProAla	486
Db	2380	ACATTTATGCTCTGAGTAATTAAGTTATCTCTGTTGGTAGTCCCACTTCTCTGCA	2439
Qy	487	CysLeuAsnProLeuLeuTyrLeuLeuPheAsnProHisPheArgAspAspLeuArgArg	506
Db	2440	TGCTCAATCCCTTCTCTACATCTTGTTCATCTCTCACTTTTAAGGAGGATCTGGTGAGC	2499
Qy	507	LeuArgProArgAla-----GlyAspSerGlyProLeuAlaTyrAlaAla	521
Db	2500	CTGAGAAAGCAAACTACGCTTGCAAGATCAAAACACCAACCAAGCTTGAATGCAATTAAC	2559
Qy	522	AlaGlyGluLeuGlySerSerCysAspSerThrGlnAlaLeuValAlaPheSerAsp	541
Db	2560	TCTGATGATGCGAAACAGCTCTCTGACTCAACTCAAGCTTGGTAAACCTTTACCAGC	2619
Qy	542	ValAspLeuIleLeuGluAlaSerGluAlaGlyArgPro-----ProGlyLeu	557
Db	2620	TCCAGCATCATTTATGACCTGCCCTCCAGTTCCGTGGCATCACCCAGCTTATCCAGTGACT	2679
Qy	558	GluThrTyrGlyPheProSerValThrLeuIleSerCys	570
Db	2680	GAGAGCTGCCATCTTCTCTCTGTCGCAATTTGTCCTCATGT	2718

RESULT 3

US-09-976-594-201
; Sequence 201, Application US/09976594
; Patent No. 6673549
; GENERAL INFORMATION:
; APPLICANT: Furness, Michael
; APPLICANT: Buchbinder, Jenny
; TITLE OF INVENTION: GENES EXPRESSED IN C3A LIVER CELL CULTURES TREATED WITH STEROIDS
; FILE REFERENCE: PA-0041 US
; CURRENT APPLICATION NUMBER: US/09/976,594
; CURRENT FILING DATE: 2001-10-12
; PRIOR APPLICATION NUMBER: 60/240,409
; PRIOR FILING DATE: 2000-10-12
; NUMBER OF SEQ ID NOS: 1143
; SOFTWARE: PERL Program
; SEQ ID NO 201

! TYPE: DNA
! ORGANISM: Homo sapiens
US-09-170-496D-277

Alignment Scores:

Pred. No.: 7 436-111 Length: 2724
Score: 1269.00 Matches: 249
Percent Similarity: 64.73% Conservative: 96
Best Local Similarity: 46.72% Mismatches: 158
Query Match: 38.37% Indels: 30
DB: 4 Gaps: 8

US-10-664-667-5 (1-633) x US-09-170-496D-277 (1-2724)

```
QY 54 LysLeuGlu**AspThrPheSerGlnLeuSerSerLeuGlnAlaLeuAspLeuSerTrp 73
DB 1162 GAAATTAAGATTGACATTTTCAGCAGATTGCTTAGCCTCGCATCGTGAATTTGCGTTGG 1221
QY 74 AnAlaIleArgSerIleHisProGluAlaPheSerThrLeuHisSerLeuValLysLeu 93
DB 1222 AACAAAATTCGCTATTATTCACCCCAATGCATTTTCACCTTTGCCATCCCTAATAAAGCTG 1281
QY 94 AspLeuThrAspAsnGlnLeuThrThrLeuProLeuAlaGlyLeuGlyLeuMetHis 113
DB 1282 GACCTATCGTCCACCTCCTCGTCTTTTCTATAACTGGGTATCATGTTAACTCAC 1341
QY 114 LeuLysLeuLysGlyAsnLeuAlaLeuSerGlnAlaPheSerLysAspSerPheProLys 133
DB 1342 TTAAATTAACAGGAATCATGCTTACAGAGCTTGATATCATCTGMAAACTTTCCAGAA 1401
QY 134 LeuArgIleLeuGluValProTyrAlaTyrGlnCysCysProTyrGlyMetCysAlaSer 153
DB 1402 CTCAGGTTATAGAAATGCTTATGCTTACCAGTGTGTCATTTGGAGTGTGAGAAT 1461
QY 154 PhePheLysAlaSerGlyClnTrp-----GluAlaGluAspLeuHis 167
DB 1462 GCCTATAAGATTTCATCAATGGAATAAAGTGACACAGCAGATGATGACGACCTTCAT 1521
QY 168 LeuAspAspGluSerSerLysArgProLeuGlyLeuLeuAlaArgGlnAlaGluAsn 187
DB 1522 -----NAGAAAGATGCTGGATGTTTCAGGCTCAGATGAA--- 1557
QY 188 HisTyrAspGlnAspLeuAspGluLeuGlnLeuMet---GluAspSerLysProHis 206
DB 1558 -----CGTGACCTTGAAGATTTCTCTGCTTGACTTTTGAGGAGAACCTGAAAGCCCTT 1608
QY 207 ProSerValGlnCysSerProThrProGlyProPheLysProCysGluTyrLeuPheGlu 226
DB 1609 CATTTCAGTGCAGTGTTCACCTTCCCGAGGCCCTTCAAAACCCCTGTGAACACCTGCTTGA 1668
QY 227 SerTrpGlyIleArgLeuAlaValTrpAlaIleValLeuSerValLeuCysAsnGly 246
DB 1669 GCCTGGCTGATCAGAAATGGAGTGGACCATAGCAGTTCTGGCACTTACTTGTGAATGCT 1728
QY 247 LeuValLeuLeuThrValPheAlaGlyGlyProAlaProLeuProProValLysPheVal 266
DB 1729 TTGTGGTACTTCAACAGTTTTCAGA---TCCCTCTGTACATTTCCGCCATTAACTGTTA 1785
QY 267 ValGlyAlaIleAlaGlyAlaAsnThrLeuThrGlyIleSerCysGlyLeuLeuAlaSer 286
DB 1786 ATTTGGGTGATCGCAGCAGTGAACATGCTCACGGGAGTCTCCAGTGGCGGTGCTGGT 1845
QY 287 ValAspAlaLeuThrPheGlyGlnPheSerGluTyrGlyAlaArgTrpGluThrGlyLeu 306
DB 1846 GTGGATGCGTTTCATTTTGGCAGCTTTGCACGACATGGTGGCTGGTGGAGAAATGGGGTT 1905
QY 307 GlyCysArgAlaThrGlyPheLeuAlaValLeuGlySerGluAlaSerValLeuLeuLeu 326
DB 1906 GTTGGCCATGTCATTTGTTTGGCCATTTTGGCTTCAGATCATCTGTTTCTGCTT 1965
QY 327 ThrLeuAlaAlaValGlnCysSerValSerValSerCysValArgAlaTyrGlyLysSer 346
DB 1966 ACTCTGGCAGCCCTGGAGCGTGGGTCTCTGTGAAATATTTCTGCAAAATTTCTGAAACGAA 2025
```

```
QY 347 ProSerLeuGlySerValArgAlaGlyValLeuGlyCysLeuAlaLeuAlaGlyLeuAla 366
DB 2026 GCTCCATTTTCTAGCCTGAAAGTAATATTGCTTCTGTGCCTCTGGCCTTGACCATG 2085
QY 367 AlaAlaLeuProLeuAlaSerValGlyGluTyrGlyAlaSerProLeuCysLeuProTyr 386
DB 2086 GCCGAGTTCCTCCCTGGTGGGAGCAAGTATGGCGCTCCCTCTCTGCTGCTGCTTGG 2145
QY 387 AlaProGluGlyGlnProAlaAlaLeuGlyPheThrValAlaLeuValMetMetAsn 406
DB 2146 -----CCTTTGGGAGCCAGCACCATGGGTACATGGTGGCTCTCTCATCTTGTCTCAAT 2199
QY 407 SerPheCysPheLeuValValAlaGlyAlaTyrIleLysLeuTyrCysAspLeuProArg 426
DB 2200 TCCCTTTGCTTCTCATGATGACCATTCCTTACACCAAGCTCTACTGCAATTTGGACAAG 2259
QY 427 GlyAspPheGluAlaValTrpAspCysAlaMetValArgHisValAlaTrpLeuIlePhe 446
DB 2260 GGAGACCTGGAGAATATTGGGACTGCTCTATGAAAAAACACATTTGCCCTGTGCTCTTC 2319
QY 447 AlaAspGlyLeuLeuTyrCysProValAlaPheLeuSerPheAlaSerMetLeuGlyLeu 466
DB 2320 ACCAACTCATCTTAACTGCCCTGTGGCTTCTTGTCTCTCTCTCTCTCTCTCTCTCT 2379
QY 467 PheProValThrProGluAlaValLysSerValLeuLeuValValLeuProLeuProAla 486
DB 2380 ACAATTTATCAGTCCCTGAAGTAATTAAGTTTATCTCTCTGGTGGTAGTCCCACTTCTGCA 2439
QY 487 CysLeuAsnProLeuLeuTyrLeuLeuPheAsnProHisPheArgAspAspLeuArg 506
DB 2440 TGTCTCAATCCCTTCTCTACATCTTGTTCATCTCTCACTTTAAGGAGGATCTGGTGAGC 2499
QY 507 LeuArgProArgAla-----GlyAspSerGlyProLeuAlaTyrAlaAla 521
DB 2500 CTGAAAGACCAACCTAGCTCTGGACAAGATCAAAACCCCAAGCTTGTATGTCAATTAAC 2559
QY 522 AlaGlyGlyLeuGluLysSerCysAspSerThrGlnAlaLeuValAlaPheSerAsp 541
DB 2560 TCTGATGATGCGAAACACAGTCTCTGCTGACTCACTCAAGCCTTGGTAACTTTTACCAGC 2619
QY 542 ValAspLeuLeuLeuGluAlaSerGluAlaGlyArgPro-----ProGlyLeu 557
DB 2620 TCCAGCATCACTTATGACCTGCCCTCCAGTTCCGTCGCCCATCACCAGCTTATCCAGTGACT 2679
QY 558 GluThrTyrGlyPheProSerValThrLeuIleSerCys 570
DB 2680 GAGAGCTGCCATCTTCTCTGTGGCATTTGTCCCATGT 2718
```

RESULT 5

US-08-866-757-1
; Sequence 1, Application US/08866757
; Patent No. 5858716
; GENERAL INFORMATION:
; APPLICANT: ELSHOURBAGY, NABIL A
; APPLICANT: LI, XIAOTONG
; APPLICANT: BERGSMAN, DEK J
; TITLE OF INVENTION: NOVEL 7TM RECEPTOR (H2CAA71)
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESS: RATNER & PRESTIA
; STREET: P.O. BOX 980
; CITY: VALLEY FORGE
; STATE: PA
; COUNTRY: USA
; ZIP: 19482
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FASTSEQ for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/866,757

; FILING DATE: 30-MAY-1997
 ; CLASSIFICATION: 536
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER:
 ; FILING DATE:
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: PRESTIA, PAUL F
 ; REGISTRATION NUMBER: 23,031
 ; REFERENCE/DOCKET NUMBER: GH-70055
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: 610-407-0700
 ; TELEFAX: 610-407-0701
 ; TELEX: 846169
 ; INFORMATION FOR SEQ ID NO: 1:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 4203 base pairs
 ; TYPE: nucleic acid
 ; STRANDEDNESS: single
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: cDNA
 ; US-08-866-757-1

Alignment Scores:
 Pred. No.: 7,256-81 Length: 4203
 Score: 958.00 Matches: 211
 Percent Similarity: 54.11% Conservative: 85
 Best Local Similarity: 38.57% Mismatches: 191
 Query Match: 28.97% Indels: 60
 DB: 2 Gaps: 9

US-10-664-667-5 (1-633) x US-08-866-757-1 (1-4203)

QY 54 LysLeuGlu***AspThrPheSerGlnLeuSerSerLeuGlnAlaLeuAspLeuSerTrp 73
 DB 697 CAAATAAGGAAGCACCCTTTCAAGGCTGTATCTCTAAGGATTTCTAGATCTCAGTAGA 756
 QY 74 AsnAlaIleArgSerIleHisProGluAlaPheSerThrLeuHisSerLeuValLysLeu 93
 DB 757 AACCTGATACATGAATTCACATGAGCTTTTGGCACACTTGGCCCATTAACCTA 816
 QY 94 AspLeuThrAspAsnGlnLeuThrThrLeuProLeuAlaGlyLeuGlyLeuMetHis 113
 DB 817 GATGTAAAGTTCAATGAATTAACCTTCTCTACGGAAGGCTGAATGGGCTAAATCAA 876
 QY 114 LeuLysLeuLysGlyAsnLeuAlaLeuSerGlnAlaPheSerLysAspSerPheProLys 133
 DB 877 CTGAAGCTGGTGGCAACTTCAAGCTGAAGAGCCTTAGCAGCAAAAGACTTTGTTAAC 936
 QY 134 LeuArgIleLeuGluValProTyrAlaTyrGlnCysProTyrGlyMetCysAlaSer 153
 DB 937 CTCAGGCTTTATCAGTACCATATGCTTATCAGTGTCTGTGCAATTTTGGGTTGTACTCT 996
 QY 154 PhePheLysAlaSerGlyGlnTrpGluAlaGluAspLeuHisLeuAspAsp----- 170
 DB 997 TAT-----GCATAATTAAACACAGAAGATAACACCTCCAGGACCACAGGTG 1044
 QY 171 ---GluGluSerSerLysArgProLeuGlyLeuLeuAlaArgGlnAlaGluAsnHisTyr 189
 DB 1045 GCACAGGAAGAAAGTACTGTCTGATGCAGCAAAATGTCACAGCACTCTTTGAAATAAGAGAA 1104
 QY 190 AspGlnAspLeuAspGluLeuGlnLeuGluMetGluAspSerLysProHisProSerVal 209
 DB 1105 CATAGTCMAATA-----ATTATC 1122
 QY 210 GlnCysSerProThrProGlyProPheLysProCysGluTyrLeuPheGluSerTrpGly 229
 DB 1123 CATTTGACACCTTCAACAGGTCCTTTAAGCCCTGTGAATATTTACTGGGAAGCTGGATG 1182
 QY 230 IleArgLeuAlaValTrpAlaIleValLeuLeuSerValLeuCysAsnGlyLeuValLeu 249
 DB 1183 ATTCGCTCTTACTGTGTGTTTCATTTCTTGTGTGCAATATTTTCAACCTGCTGTGTATT 1242
 QY 250 LeuThrValPheAlaGlyGlyProAlaProLeuProValLysPheValValGlyAla 269

DB 1243 TTAACAACATTTGTCATCTTGTACATCA---CTGCCTTCGTCCAAATTTGTTTATAGGCTTG 1299
 QY 270 IleAlaGlyAlaAsnThrLeuThrGlyLysSerCysGlyLeuLeuAlaSerValAspAla 289
 DB 1300 ATTTCTGTGTAACTATTATCATGGGAATCTATACTGGCATCTCTAACTTTTCTTGATGCT 1359
 QY 290 LeuThrPheGlyGlnPheSerGluTyrGlyAlaArgTrpGluThrGlyLeuGlyCysArg 309
 DB 1360 GTGTCCTGGGCGAGATTCGCTGAATTTGGCATTTGGTGGGAAACTGGCAGTGGGTGCAAA 1419
 QY 310 AlaThrGlyPheLeuAlaValLeuGlySerGluAlaSerValLeuLeuLeuThrLeuAla 329
 DB 1420 GTRAACTGGGTTTCTTCAGTTTTCTCTCAGAAAGTCCCATATTTTATTATTAATGCTAGCA 1479
 QY 330 AlaValGlnCysSerValSerCysValArgAlaTyrGlyLysSerProSerLeu 349
 DB 1480 ACTGTCGAAAGAAAGCTTATCTGCAAAAGATATATGAAATAATGGGAAGAGCAATCATCTC 1539
 QY 350 GlySerValArgAlaGlyValLeuGlyCysLeuAlaLeuAlaGlyLeuAlaAlaLeu 369
 DB 1540 AACAGTTCCGGTTCGCTTCCTAGTCTTCTAGTGTACAGTAACAGGCTGTTTT 1599
 QY 370 ProLeuAlaSerValGlyGluTyrGlyAlaSerProLeuCysLeuProTyrAlaProPro 389
 DB 1600 CCCCTTTTCCATAGAGGGGAATATCTGCATCACCCTTTGTTGGCCATTT-----CCT 1653
 QY 390 GluGlnProAlaAlaLeuGlyPheThrValAlaLeuValMetMetAsnSerPheCys 409
 DB 1654 ACAGGTGAACCGCATCATTTAGGATTCACGTGTAACGTAGTGTATTAAACTACTAGCA 1713
 QY 410 PheLeuValValAlaGlyAlaTyrIleLysLeuTyrCysAspLeuProArgGlyAspPhe 429
 DB 1714 TTTTATTAAAGCCGTTATCTACATAAGTATCTGCAACTTGGAAAGAGAGACCTC 1773
 QY 430 GluAlaValTrpAspCysAlaMetValArgHisValAlaTrpLeuIlePheAlaAspGly 449
 DB 1774 TCAGAAAACCTCACAACTAGCATGATTAAAGCATGCTGCTGGCTAATCTTCACCAATTC 1833
 QY 450 LeuLeuTyrCysProValAlaPheLeuSerPheAlaSerMetLeuGlyLeuPheProVal 469
 DB 1834 ATCTTTTCTCCCTGTGGGGTTTTTTTTCATTTTGCACTTGTATCTCTATC 1893
 QY 470 ThrProGluAlaValLysSerValLeuValValLeuProLeuProAlaCysLeuAsn 489
 DB 1894 AGCCCCGAAATAAAGAGTCTGTACTGTATATTTTCCATTTGCTCTGCTGCTGAAT 1953
 QY 490 ProLeuLeuTyrLeuLeuPheAsnProHisPheArgAspAsp-----LeuArg 505
 DB 1954 CCAGTCTCTGTATGTTTCTTCAACCCAAAGTTTAAAGAGACTTGAAGTTACTGAAGCA 2013
 QY 506 ArgLeuArgProArgAlaGlyAspSerGlyProLeuAlaTyrAlaAlaGlyGluLeu 525
 DB 2014 CGTGTACCAAGAAAGTGGATCAGTTTTCAGTTTCCATCAGTAGCAAGGTGGTGTCTG 2073
 QY 526 GluLys-----SerSer 529
 DB 2074 GAACAGGATTTCTACTACGACTGTGGCATGTACTCATTTGCGAGGCAACCTGACTGTT 2133
 QY 530 CysAspSerThrGlnAlaLeuValAlaPheSerAspValAsp-----LeuIleLeu 546
 DB 2134 TGGGACTGTGCGAATCGTTTCTTTAAACAAGCCAGTATCATGCAACACTTGATA--- 2190
 QY 547 GluAlaSerGluAlaGlyArgProGlyLeuGluThrTyrGlyPheProSerValThr 566
 DB 2191 -----AAATCACACACAGCTGCTCTGCTGATGGCA 2217
 QY 567 LeuIleSerCysGlnGlnPro 573
 DB 2218 GTGGCTCTTGGCAAGAGCT 2238
 RESULT 6
 US-09-153-593-1


```
; GENERAL INFORMATION:
; APPLICANT: Roopa, Reddy
; APPLICANT: Guegler, Karl, J.
; APPLICANT: Au-Young, Janice
; TITLE OF INVENTION: COMPOSITION FOR DETECTION OF GENES ENCODING MEMBRANE-ASSOCIATED P
; FILE REFERENCE: PA-0013 US
; CURRENT APPLICATION NUMBER: US/09/495,050A
; CURRENT FILING DATE: 2000-01-31
; PRIOR APPLICATION NUMBER: 60/118,318
; PRIOR FILING DATE: February 1, 1999
; NUMBER OF SEQ ID NOS: 305
; SOFTWARE: PERL Program
; SEQ ID NO 214
; LENGTH: 2612
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No. 6492505 2470285CT1
US-09-495-050A-214

Alignment Scores:
Pred. No.: 4,32e-81 Length: 2612
Score: 957.00 Matches: 211
Percent Similarity: 54.11% Conservative: 85
Best Local Similarity: 38.57% Mismatches: 191
Query Match: 28.94% Indels: 60
DB: 4 Gaps: 9

US-10-664-667-5 (1-633) x US-09-495-050A-214 (1-2612)

Qy 54 LysLeuGlu***AspThrPheSerGlnLeuSerSerLeuGlnAlaLeuAspLeuSerTrp 73
Db 806 CAAATAAAGGAAGCACCCTTTCAAGGCTGTATATCTCTAAGGATTCAGATCTGAGTAGA 865
Qy 74 AsnAlaIleArgSerIleHisProGluAlaPheSerThrLeuHisSerLeuValLysLeu 93
Db 866 AACCTGATACATGAATTCACAGTAGAGCTTTGGCCACACTTGGGCCAATAACCTA 925
Qy 94 AspLeuThrAspAsnGlnLeuThrLeuProLeuAlaGlyLeuGlyLeuMetHis 113
Db 926 GATGTAAGTTTCAATGAATTAACCTTCCTTCTACGGAAGGCTGAAATGGGCTAAATCAA 985
Qy 114 LeuLysLeuLysGlyAsnLeuAlaLeuSerGlnAlaPheSerLysAspSerPheProLys 133
Db 986 CTGAACACTTGGGCAACTTCAAGCTGAAGAAGCCCTTAGCAGCAAAAGACTTTGTTAAC 1045
Qy 134 LeuArgIleLeuGluValProTyrAlaTyrGlnCysCysProTyrGlyMetCysAlaSer 153
Db 1046 CTCAGGCTTTATCAGTAGACCATATGCTTATCAGTGCTGTGCAATTTTGGGGTTGTGACTCT 1105
Qy 154 PhePheLysAlaSerGlyGlnTrpGluAlaGluAspLeuHisLeuAspAsp----- 170
Db 1106 TAT-----GCAAAATTTAAACACAGAAATAAACAGCTCCAGGACCACAGGTGTG 1153
Qy 171 ---GluGluSerSerLysArgProLeuGlyLeuLeuAlaArgGlnAlaGluAsnHisTyr 189
Db 1154 GCACAGGAGAAAGGTACTGTGTAGTGCAGCAATGTCCAGACACTCTTGAATAATGAAGAA 1213
Qy 190 AspGlnAspLeuAspGluLeuGlnLeuMetGluAspSerLysProHisProSerVal 209
Db 1214 CATAGTCAATA-----ATTATC 1231
Qy 210 GlnCysSerProThrProGlyProPheLysProCysGluTyrLeuPheGluSerTrpGly 229
Db 1232 CATGTACACCTTCAACAGAGTGCTTTTAAAGCCCTGTGAATATTATTACTGGGAAGCTGGATG 1291
Qy 230 IleArgLeuAlaValTrpAlaIleValLeuLeuSerValLeuCysAsnGlyLeuValLeu 249
Db 1292 ATTGCTCTTACTGTGTGGTTCATTTCTTGGTCATTATTTTCAACCTGCTGTGTTATT 1351
Qy 250 LeuThrValPheAlaGlyGlyProAlaProLeuProValLysPheValValGlyAla 269
Db 1351
```

```
Db 1352 TTAACAACAATTCATCTTGTACATCA--CTGCCTTCGTCACAAATTTGTTATAGGCTTG 1408
Qy 270 IleAlaGlyAlaAsnThrLeuThrGlyIleSerCysGlyLeuLeuAlaSerValAspAla 289
Db 1409 ATTTCGTGTCTAATTAATTCATGGGAATCTATCTGGCATCTTAACCTTTCTTGATGCT 1468
Qy 290 LeuThrPheGlyGlnPheSerGluTyrGlyAlaArgTrpGluThrGlyLeuGlyCysArg 309
Db 1469 GTGTCTCTGGGCGAGATTCGCTGAATTTGGCATTTGGTGGGAAACTGGCAGTGGCTGCAA 1528
Qy 310 AlaThrGlyPheLeuAlaValLeuGlySerGluAlaSerValLeuLeuLeuThrLeuAla 329
Db 1529 GTAGCTGGGTTTCTTTCAGATTTTCTCTCAGAAAGTGCCATATTTTATTAATGCTAGCA 1588
Qy 330 AlaValGlnCysSerValSerValSerCysValArgAlaTyrGlyLysSerProSerLeu 349
Db 1589 ACTGTCGAAAGAGCTTATCTGCAAAAGATATTAATGAAATGGGAAGCAATCATCTC 1648
Qy 350 GlySerValArgAlaGlyValLeuGlyCysLeuAlaLeuAlaGlyLeuAlaAlaLeu 369
Db 1649 AACAGTTCGGGTGCTGCCCTTTTGGCTTCTTAGTGTCTACAGTAGCAGGCTGTTTT 1708
Qy 370 ProLeuAlaSerValGlyGluTyrGlyAlaSerProLeuCysLeuProTyrAlaProPro 389
Db 1709 CCCCTTTTCCATAGAGGGGAATATCTGCATCACCCCTTTGTTGGCAATTT-----CCT 1762
Qy 390 GluGlyGlnProAlaAlaLeuGlyPheThrValAlaLeuValMetMetAsnSerPheCys 409
Db 1763 ACAGGTGAAGAGCCATCATTAGGATTCACGTGAACGTAGTGTATTAATAACTACTAGCA 1822
Qy 410 PheLeuValValAlaGlyAlaTyrIleLysLeuTyrCysAspLeuProArgGlyAspPhe 429
Db 1823 TTTTATTAATGCGCTTATCTACACTAGCTATCTGCAACTTGGAAAGAGGAGCCTC 1882
Qy 430 GluAlaValTrpAspCysAlaMetValArgHisValAlaTrpLeuIlePheAlaAspGly 449
Db 1883 TCAGAAAACCTCACAATCTAGCATGATTAAAGCATGCTGGTGGTAAATCTTCCACCAATGC 1942
Qy 450 LeuLeuTyrCysProValAlaPheLeuSerPheAlaSerMetLeuGlyLeuPheProVal 469
Db 1943 ATCTTTTTCGCTGCTGGCGTGTTCATTTGACCATTTGATCATCTGCAATCTCTATC 2002
Qy 470 ThrProGluAlaValLysSerValLeuLeuValLeuProLeuProAlaCysLeuAsn 489
Db 2003 AGCCCGGAATATGAAGTCTGTACTCTGATATTTTTCATGCTGCTGCTGCTGCTGAT 2062
Qy 490 ProLeuLeuTyrLeuLeuPheAsnProHisPheArgAspAsp-----LeuArg 505
Db 2063 CCAGTCTGTATGTTTCTTCAACCCCAAGTTTAAAGAGAGACTGGAAGTTTACTGAAGCGA 2122
Qy 506 ArgLeuArgProArgAlaGlyAspSerGlyProLeuAlaTyrAlaAlaAlaGlyGluLeu 525
Db 2123 CGTGTTCACCAAGAAAAGTGGATCAGTTTTCATCTCCATCAGTAGCCAAAGGTGTTGCTG 2182
Qy 526 GluLys-----SerSer 529
Db 2183 GACAGGATTTCTACTACGACTGTGGCATGTACTCACAATTTGCGGCAACCTGACTGTT 2242
Qy 530 CysAspSerThrGlnAlaLeuValAlaPheSerAspValAsp-----LeuIleLeu 546
Db 2243 TGGGACTGTGCGAATCGTTTCTTTAAACAAAGCCAGTATCATGCAAAACACTTGATA-- 2299
Qy 547 GluAlaSerGluAlaGlyArgProGlyLeuGluThrTyrGlyPheProSerValThr 566
Db 2300 -----AAATCACACAGCTCTCTCGATTCGGC 2326
Qy 567 LeuIleSerCysGlnGlnPro 573
Db 2327 GTGGCTTCTTGCCAAAGACCT 2347

RESULT 8
US-09-397-787-240
; Sequence 240, Application US/09397787
```

```
; Patent No. 6468758
; GENERAL INFORMATION:
; APPLICANT: Benson, Darin R.
; APPLICANT: Lodes, Michael J.
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: King, Gordon E.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR OVARIAN
; TITLE OF INVENTION: CANCER THERAPY AND DIAGNOSIS
; FILE REFERENCE: 210121.466C2
; CURRENT APPLICATION NUMBER: US/09/397,787
; CURRENT FILING DATE: 1999-09-16
; NUMBER OF SEQ ID NOS: 334
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 240
; LENGTH: 453
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-397-787-240

Alignment Scores:
Pred. No.: 2,3e-63 Length: 453
Score: 759.00 Matches: 149
Percent Similarity: 98.68% Conservative: 0
Best Local Similarity: 98.68% Mismatches: 1
Query Match: 22.95% Indels: 1
DB: 3 Gaps: 0

US-10-664-667-5 (1-633) x US-09-397-787-240 (1-453)
Qy 444 LeullePheAlaAaspGlyLeuLeuTyrcysProValAlaPheLeuSerPheAlaSerMet 463
Db 3 CTCATCTTCGACAGCGGCTCTCTACTGTGCTCCGTCGCGCTTCTCAGCTTCGCTCCATG 62
Qy 464 LeuGlyLeuPheProValThrProGluAlaValLysSerValLeuValValLeuPro 483
Db 63 CTGGGCTCTTCCCTGTCTCAGCCCGGAGCGGCTCAAGTCTGCTCTCTGCTGTGCTGCC 122
Qy 484 LeuProAlaCysLeuAsnProLeuLeuTyrcysProValAlaPheLeuSerPheAlaSerMet 503
Db 123 CTGCTGCTGCTTCAACCCCTGCTGCTGCTGCTTCAACCCCTGCTGCTGCTGCTGCTG 182
Qy 504 LeuArgArgLeuArgProArgAlaGlyAaspSerGlyProLeuAlaTyrcysAlaAlaGly 523
Db 183 CTTCGCGGCTTCGCGCCCGGAGGAGCTCAGGCGGCTGAGCTGCTGCTGCTGCTGCTG 242
Qy 524 GluLeuGluLysSerSerCysAaspSerThrGlnAlaLeuValAlaPheSerAaspValasp 543
Db 243 GAGCTGGAGAGAGCTTCCCGTATTTCTACCCAGGCGGCTGGTAGCTTCTCTGATGTGGAT 302
Qy 544 LeulleLeuGluAlaSerGluAlaGlyArgProProGlyLeuGluThrTyrcysGlyPhePro 563
Db 303 CTCATCTTGGAGCTTCTGAGCTGGGCGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTG 362
Qy 564 SerValThrLeuLysSerCysGlnGlnProGlyAlaProArgLeuGluGly-SerHisCy 583
Db 363 TCAGTGACCTCATCTCTCTGTCAGCAGCAGGCGGCGGCTGAGGCTGAGGCGGCAAGCATTG 422
Qy 583 aValGluProGluGlyAenHisPheGlyAen 593
Db 423 TGTAGACCCAGAGGGGAACCACTTTGGGAAC 453

RESULT 9
US-09-495-050A-220
; Sequence 220, Application US/09495050A
; Patent No. 6492505
; GENERAL INFORMATION:
; APPLICANT: Roopa, Reddy
; APPLICANT: Guegler, Karl, J.
; APPLICANT: Au-Young, Janice
; TITLE OF INVENTION: COMPOSITION FOR DETECTION OF GENES ENCODING MEMBRANE-ASSOCIATED P
; FILE REFERENCE: PA-0013 US
; CURRENT APPLICATION NUMBER: US/09/495,050A
; CURRENT FILING DATE: 2000-01-31
```

```
; PRIOR APPLICATION NUMBER: 60/118,318
; PRIOR FILING DATE: February 1, 1999
; NUMBER OF SEQ ID NOS: 305
; SOFTWARE: PERL Program
; SEQ ID NO 220
; LENGTH: 723
; TYPE: DNA
; ORGANISM: Homo sapiens
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No. 6492505 2603450CT1
US-09-495-050A-220
```

```
Alignment Scores:
Pred. No.: 1.95e-47 Length: 723
Score: 595.50 Matches: 116
Percent Similarity: 67.36% Conservative: 47
Best Local Similarity: 47.93% Mismatches: 76
Query Match: 18.01% Indels: 4
DB: 4 Gaps: 2
```

US-10-664-667-5 (1-633) x US-09-495-050A-220 (1-723)

```
Qy 234 ValTrpAlaIleValLeuLeuSerValLeuCysAsnGlyLeuValLeuLeuThrValPhe 253
Db 6 GTGTGGACCATAGCAGTTCTGGCATTCTTGTATGCTTTGGTCACTTCAACAGTTTTTC 65
Qy 254 AlaGlyGlyProAlaProLeuProValLysPheValValGlyAlaIleAlaGlyAla 273
Db 66 AGA---TCCCTCTGTACATTTCCCTCATTAACCTGTTAATGGGTCATTCGACAGTG 122
Qy 274 AsnThrLeuThrGlyIleSerCysGlyLeuLeuAlaSerValAspAlaLeuThrPheGly 293
Db 123 AACATGCTCAGCGGAGTCTCCAGTCCGCTGCTGGTGTGGATGCGTTTCACTTTTGGC 182
Qy 294 GlnPheSerGluTyrcysAlaArgTrpGluThrGlyLeuGlyCysArgAlaThrGlyPhe 313
Db 183 AGCTTTGACACACATGTCCTGCTGGGAGAAATGGGTTGGTTCATTCATTTGGTTT 242
Qy 314 LeuAlaValLeuGlySerGluAlaSerValLeuLeuLeuThrLeuAlaAlaValGlnCys 333
Db 243 TTGTCATTTTGGCTTTCAGAAATCATCTGTTTCTGCTTACTCTGGCAGCCCTGGAGGCT 302
Qy 334 SerValSerValSerCysValArgAlaTyrcysGlyLysSerProSerLeuGlySerValArg 353
Db 303 GGGTTCTCTGTGAATATTCTGCAAAATTTGAAACGAAAGCTCCATTTTCTAGCCTGAAA 362
Qy 354 AlaGlyValLeuGlyCysLeuAlaLeuAlaGlyLeuAlaAlaLeuProLeuAlaSer 373
Db 363 GTAATCATTTTGTCTCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 422
Qy 374 ValGlyGluTyrcysAlaSerProLeuCysLeuProTyrcysAlaProGluGlnPro 393
Db 423 GGCAGAGAGTATGGCGCTCCCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 476
Qy 394 AlaAlaLeuGlyPheThrValAlaLeuValMetMetAsnSerPheCysPheLeuVal 413
Db 477 AGCACCATGGGCTACATGTCGCTCATCTTGTCTCAATTCCTTTGCTTCTCTCATGATG 536
Qy 414 AlaGlyAlaTyrcysLeuTyrcysAspLeuProArgGlyAspPheGluAlaValTrp 433
Db 537 ACCATTGCTACACCAAGCTCTACTGCAATTTGGACAGGGAGACCTCGAGAAATATTTGG 596
Qy 434 AspCysAlaMetValArgHisValAlaTrpLeuIlePheAlaAaspGlyLeuLeuTyrcys 453
Db 597 GACTGCTCTATGGTAAACACATTTGCCCTGTGCTTCTTACCAACCTGCTTCAACTGC 656
Qy 454 ProValAlaPheLeuSerPheAlaSerMetLeuGlyLeuPheProValThrProGluAla 473
Db 657 CTTGTGGCTTCTTGTCTCTCTCTCTTATATAACCTTATCATC-ATCAGTCTCTGAAGTA 715
Qy 474 ValLys 475
:::|
```

Db 716 ATTAAG 721

RESULT 10

US-08-487-886-1

; Sequence 1, Application US/08487886

; Patent No. 5744448

; GENERAL INFORMATION:

; APPLICANT: Kelton, Christie Ann

; APPLICANT: Schweickhardt, Rene Lynn

; APPLICANT: Cheng, Shirley Vui Yen

; APPLICANT: Nugent, No. 57444888 Patrice

; TITLE OF INVENTION: Human Follicle Stimulating

; TITLE OF INVENTION: Hormone Receptor

; NUMBER OF SEQUENCES: 2

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Stephan P. Williams,

; ADDRESSEE: Ares-Serono, Inc.

; STREET: Exchange Place, 37th floor

; CITY: Boston

; STATE: MA

; COUNTRY: USA

; ZIP: 02109

; COMPUTER READABLE FORM:

; MEDIUM TYPE: 3.5" diskette, 1.44 MB, high density

; COMPUTER: IBM PS/2, model 55 SX

; OPERATING SYSTEM: MS-DOS version 4.0

; SOFTWARE: VAX/VMS Massll via Kermit to IBM MS-DOS

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/487,886

; FILING DATE: 07-JUN-1995

; CLASSIFICATION: 514

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: 07/670,085

; FILING DATE: 15-MAR-1991

; ATTORNEY/AGENT INFORMATION:

; NAME: Williams, Stephan P.

; REGISTRATION NUMBER: 28546

; REFERENCE/DOCKET NUMBER: US/252

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (617) 723-1300

; TELEFAX: (617) 723-8923

; INFORMATION FOR SEQ ID NO: 1:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 2179

; TYPE: Nucleic acid

; STRANDEDNESS: Double

; TOPOLOGY: Linear

; MOLECULE TYPE: cDNA to mRNA

; ORIGINAL SOURCE:

; ORGANISM: Homo sapiens

; TISSUE TYPE: Testis

; IMMEDIATE SOURCE:

; LIBRARY: lgt11 cDNA library, Clontech #HL1010b

; CLONE: pFHSR11-11, pFHSR15-6

; FEATURE:

; NAME/KEY: protein coding region

; LOCATION: 75 to 2159

US-08-487-886-1

Alignment Scores:

Pred. No.:	2,08e-29	Length:	2179
Score:	414.50	Matches:	131
Percent Similarity:	42.94%	Conservative:	94
Best Local Similarity:	25.00%	Mismatches:	232
Query Match:	12.53%	Indels:	67
DB:	1	Gaps:	13

US-10-664-667-5 (1-633) x US-08-487-886-1 (1-2179)

Qy 55 LeuGlu***AspThrPheSerGlnLeuSer---SerLeuGlnAlaLeuAspLeuSerTrp 73

Db 552 ATTGAAGAAATCTTTTCGTGGGGCTGAGCTTGGAAAGTGTGATTCTATGGCTGAATAAG 611

Qy 74 AsnAlaIleArgSerIleHisProGluAlaPhe----- 84

Db 612 AATGGGATTCAAGAAATACACAACACTGTGCATTCAATGGAACCACTAGATGAGCTGAAT 671

Qy 85 -----SerThrLeuHisSerLeu----- 90

Db 672 CTAAGCGATAATAATATTAGAGAAATTGCCCTAATGATGTTTCCACGAGGCTCTGGA 731

Qy 91 ---ValLeuLeuAspLeuThrAspAsnGlnLeuThrLeuProLeuAlaGlyLeuGly 109

Db 732 CCACTCATCTTAGATATTTCAAGAACAGGATCCATTCCTGCTAGCTATGGTTAGAA 791

Qy 110 GlyLeuMetHisLeuLeuLysGlyAsnLeuAlaLeuSerGlnAlaPheSerLysAsp 129

Db 792 AATCTTAAGAAGCTGAGGGCCAGTCCGACTTACAACCTTAAAAAAGCTCTACTCTGAA 851

Qy 130 SerPheProLysLeuArgIleLeuGluValProTyrAlaTyrGlnCysCysProTyrGly 149

Db 852 AAGCTTGTGCGCCCTCATGGAAGCCAGCTCACCTATCCAGCCATCTGCTGCTTT--- 908

Qy 150 MetCysAlaSerPhePheLysAlaSerGlyGlnTrpGluAlaGluAspLeuHis----- 167

Db 909 -----GCAACTGGAGAGCGCAATCTCTGAGCTTCATCCATT 947

Qy 168 -----LeuAspAspGluGluSerSerLysArgProLeu 178

Db 948 TGCAACAAATCTATTTTAAGGCAAGAAGTTGATTATATGACTCAGACTAGGGGTCAGAGA 1007

Qy 179 GlyLeuLeuAlaArgGlnAlaGluAsnHisTyrAspGlnAspLeuAspGluLeuGlnLeu 198

Db 1008 TCCTCTCTGGCAGAAGACAATGATCCAGCTACAGCAGAGGATTTGACATGACGTACACT 1067

Qy 199 GluMetGluAspSerLysProHisPro-----SerValGlnCysSerProThrProGly 216

Db 1068 GAGTTTGACTATGACTTATGCAATGAAGTGTGAGTGAGTGCTGCTCCCTTAAGCCAGAT 1127

Qy 217 ProPheLysProCysGluTyrLeuPheGluSerTrpGlyIleArgLeuAlaValTrpAla 236

Db 1128 GCATTCAACCCATGTGAAGATATCATGGGGTACAACTCCTCAGAGTCTGATATGTTT 1187

Qy 237 IleValLeuLeuSerValLeuCysAsnGlyLeuValLeuLeuThrValPheAlaGlyGly 256

Db 1188 ATCAGCATCTCGCCCATCTACTGGGAAC---ATCATAGTGTAGTAGTCTTAATACACGC 1244

Qy 257 ProAlaProLeuProValLysPheValValGlyAlaIleAlaGlyAlaAsnThrLeu 276

Db 1245 CAATATAACTCAGTCCCGAGTTCTTATGTGCACTGCTGCTTGTGTGATCTCTGC 1304

Qy 277 ThrGlyIleSerCysGlyLeuLeuAlaSerValAspAlaLeuThrPheGlyGlnPheSer 296

Db 1305 ATTGGAATCTACCTGCTGCTCATTTGCATCATGATATCCATCCAGAGACCAATATCAC 1364

Qy 297 GluTyrGlyAlaArgTrpGluThrGlyLeuGlyCysArgAlaThrGlyPheLeuAlaVal 316

Db 1365 AACTATGCCATTGACTGGCAAACTGGGGCAGGCTGTGATGCTGTGCTTTTTCACCTGTC 1424

Qy 317 LeuGlySerGluAlaSerValLeuLeuLeuThrLeuAlaAlaValGln----- 332

Db 1425 TTGCGCAGTGAGCTGTGAGTCTACCTCTGACAGCTATCATCTGGAAAGATGGCATACC 1484

Qy 333 -----CysSerValSerValSerCysValArgAlaTyrGlyLysSerProSerLeuGly 350

Db 1485 ATCAGCGATCCCATGCGAGCTGGAGTGC-----AAGGTGACGCTCGGCCAT 1529

Qy 351 SerValArgAlaGlyValLeuGlyCysLeuAlaLeuAlaGlyLeuAlaAlaLeuPro 370

Db 1530 GCTGCCAGTGTGATGTTGGTGGCTGGATT---TTTGTCTTTTGCAGCTGCGCTTTTCCC 1586

Qy 371 LeuAlaSerValGlyGluTyrGlyAlaSerProLeuCysLeuProTyrAlaProProGlu 390

Db 1587 ATCTTTGGCATGACGAGCTACATGAGGTGAGCTGCTGCTGCCATG-----GATATT 1640

Qy 391 GlyGlnProAlaAlaLeuGlyPheThrValAlaLeuValMetMetAsnSerPheCysPhe 410

```

Db 1641 GACAGCCCTTGTACAGCTATGTATGCTCCCTCTGTGCTCAATGTCTCGCCCTTT 1700
Qy 411 LeuValAlaGlyAlaTyrIleLysLeuTyrCysAspLeuProArgGlyAspPhe--- 429
Db 1701 GTGGTCATCTGGCTGTATATACATCTACCTACAGTGGCGAAGCCCAACATCGTG 1760
Qy 430 GluAlaValTyrAspCysAlaMetValArgHisValAlaTyrLeuIlePheAlaAspGly 449
Db 1761 TCCTCCTCTAGTGACACAGGATGCCAAGCGCATGGCGCATGCTCATCTTCACTGATTC 1820
Qy 450 LeuLeuTyrCysProValAlaPheLeuSerPheAlaSerMetLeuGlyLeuPheProVal 469
Db 1821 CTCTGCAATGGACCCATTTCTTCTTCCCAATTTCTGCCCTCCCTCAAGGTGCCCTCATC 1880
Qy 470 ThrProGluAlaValLysSerValLeuValValLeuProLeuProAlaCysLeuAsn 489
Db 1881 ACTGTGTCACCAAGCAAGATTCGTGCTGTCTGTGTTTACCCCATCACTCTCTGTCAC 1940
Qy 490 ProLeuLeuTyrLeuLeuPheAsnProHisPheArgAspLeuArgArgLeuArgPro 509
Db 1941 CCTTCTCTATGCCATCTTTACCAAAACCTTCGACAGATTTCTTCACTTCTGTGAGC 2000
Qy 510 ArgAlaGlyAspSerGlyProLeuAlaTyrAlaAlaGlyGluLeuGluLysSerSer 529
Db 2001 AAGTGTGGC-----TGCTATGAATGCAAGGCCCAAAATTTATAGGACAGAA 2045
Qy 530 CysAspSerThr 533
Db 2046 ACTTCATCCACT 2057

```

RESULT 11

```

US-08-531-070A-1
; Sequence 1, Application US/08531070A
; Patent No. 5851768
; GENERAL INFORMATION:
; APPLICANT: de la Chapelle, Albert
; APPLICANT: Aittomaki, Kristiina
; APPLICANT: Huhtaniemi, Ilpo
; TITLE OF INVENTION: Method For Diagnosis Of Ovarian Dysgenesis
; NUMBER OF SEQUENCES: 18
; CORRESPONDENCE ADDRESS:
; ADDRESS: Marshall, O'Toole, Gerstein, Murray & Borun
; STREET: 6300 Sears Tower, 233 South Wacker Drive
; CITY: Chicago
; STATE: Illinois
; COUNTRY: United States of America
; ZIP: 60606-6402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/531.070A
; FILING DATE: 20-SEP-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Gass, David A.
; REGISTRATION NUMBER: 38,153
; REFERENCE/DOCKET NUMBER: 28113/32879
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312/474-6300
; TELEFAX: 312/474-0448
; TELEX: 25-3856
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2179 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
US-08-531-070A-1

```

```

Alignment Scores:
Pred. NO.: 2.08e-29 Length: 2179
Score: 414.50 Matches: 131
Percent Similarity: 42.94% Conservative: 94
Best Local Similarity: 25.00% Mismatches: 232
Query Match: 12.53% Indels: 67
DB: 2 Gaps: 13
US-10-664-667-5 (1-633) x US-08-531-070A-1 (1-2179)
Qy 55 LeuGlu***AspThrPheSerGlnLeuSer---SerLeuGlnAlaLeuAspLeuSerTrp 73
Db 552 ATTGAAGAATAATCTTTCTGGGGCTGAGCTTTGAAAGTGTGATTTCTATGCTGAATAAG 611
Qy 74 AsnAlaIleArgSerIleHisProGluAlaPhe----- 84
Db 612 AATGGGATTCAAGAAATACACAACTGTGCATTCATGGAACCAACTAGATGAGTGAAT 671
Qy 85 -----SerThrLeuHisSerLeu----- 90
Db 672 CTAAGCGATAATAATAATTTAGAAGAAATGCTTAATGATGTTTCCACGGAGCCTCTGGA 731
Qy 91 ---ValLysLeuAspLeuThrAspAsnGlnLeuThrLeuProLeuAlaGlyLeuGly 109
Db 732 CCAGTCAATCTAGATATTTTCAAGAACCAAGGATCCATTCCTCTGCTAGCTATGGCTTAGAA 791
Qy 110 GlyLeuMetHisLeuLysLeuLysGlyAsnLeuAlaLeuSerGlnAlaPheSerLysAsp 129
Db 792 AATCTTAAGAAGCTGAGGGCCAGCTGACTTACAACTTAAAAAGCTGCCTACTCTGGA 851
Qy 130 SerPheProLysLeuArgIleLeuGluValProTyrAlaTyrGlnCysCysProTyrGly 149
Db 852 AAGCTTGTGGCCCTCATGGAAGCCAGCTACCTATCCAGCCATTTGCTGTGCTTT--- 908
Qy 150 MetCysAlaSerPhePheLysAlaSerGlyGlnTrpGluAlaGluAspLeuHis----- 167
Db 909 -----GCAAACTGGAGAGCGCAAAATCTCTGAGCTTCATCCAAAT 947
Qy 168 -----LeuAspAspGluGluSerSerLysArgProLeu 178
Db 948 TGCACAAATCTATTTAAGCAAGAAAGTTGATTTATATGACTCAGACTAGGGGTGAGAGA 1007
Qy 179 GlyLeuLeuAlaArgGlnAlaGluAsnHisTyrAspGlnAspLeuAspGluLeuGlnLeu 198
Db 1008 TCCTCTCTGGCAGAAACAATGAGTCCAGCTACAGCAGAGGATTTGACATGACGTACACT 1067
Qy 199 GluMetGluAspSerLysProHisPro-----SerValGlnCysSerProThrProGly 216
Db 1068 GAGTTTGACTATGACTTATGCAATGAAGTGGTTGACGTGACCTGCTCCCTCAAGCCAGAT 1127
Qy 217 ProPheLysProCysGluTyrLeuPheGluSerTrpGlyIleArgLeuAlaValTrpAla 236
Db 1128 GCATTCACCAACCATGTGAGATATCATGGGGTACAACTCCCTCAGAGTCTGATATGGTTT 1187
Qy 237 IleValLeuLeuSerValLeuCysAsnGlyLeuValLeuLeuThrValPheAlaGlyGly 256
Db 1188 ATCAGCATCTCTGGCCATCCTGAGGAACT---ATCATAGTGTAGTGTCTTAACCTACAC 1244
Qy 257 ProAlaProLeuProProValLysPheValValGlyAlaIleAlaGlyAlaAsnThrLeu 276
Db 1245 CAATATAAATCTCACAGTCCCGAGGTTCCTTATGTGCAACCTGGGCTTTTGTGATCTCTGC 1304
Qy 277 ThrGlyIleSerCysGlyLeuLeuAlaSerValAspAlaLeuThrPheGlyGlnPheSer 296
Db 1305 ATTGGAATCTACCTGCTGCTCATTCATGATCATGATATATCCATACCAAGAGCAATATCAC 1364
Qy 297 GluTyrGlyAlaArgTrpGluThrGlyLeuGlyCysArgAlaThrGlyPheLeuAlaVal 316
Db 1365 AACTATGCCATTGACTGGCAAACTGGGCGAGGCTGTGATGCTGTGCTGCTTTTTCACCTGTC 1424
Qy 317 LeuGlySerGluAlaSerValLeuLeuLeuThrLeuAlaValGln----- 332

```


STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
LIBRARY: GENBANK
CLONE: G182770
US-09-016-434-1209

Alignment Scores:

Pred. No.: 2,4e-29 Length: 2393
Score: 414.50 Matches: 123
Percent Similarity: 44.20% Conservative: 94
Best Local Similarity: 25.05% Mismatches: 231
Query Match: 12.53% Indels: 43
DB: 4 Gaps: 11

US-10-664-667-5 (1-633) x US-09-016-434-1209 (1-2393)

```
Qy 64 SerSerLeuGlnAlaLeuAspLeuSer---TrpAlaAlaIleArgSerIleHisProGlu 82
Db 643 ACCCAACTAGATGCGAGTGAATCTAAGCGATAATAATTTAGAGAATTCGCTAATGAT 702
Qy 83 AlaPheSerThrLeuHisSerLeuValLysLeuAspLeuThrAspAsnGlnLeuThrThr 102
Db 703 GTTTTCCAGGAGCCTCTGGACCACTGATTTAGATATTTCAAGAACAGGATCCATCCC 762
Qy 103 LeuProLeuAlaGlyLeuGlyLeuMethHisLeuLysLeuLysGlyAsnLeuAlaLeu 122
Db 763 CTGCTAGCTATGGCTTAGAAAATCTTAAGAAGCTGAGGGCCAGGTCCAGCTTACAACITTA 822
Qy 123 SerGlnAlaPheSerLysAspSerPheProLysLeuArgIleLeuGluValProTyrAla 142
Db 823 AAAAAGTCGCTACTCTCGAAAAGCTTGTGCGCCTCATGGAAGCCAGCCTCACCTATCCC 882
Qy 143 TyrGlnCysCysProTyrGlyMetCysAlaSerPhePheLysAlaSerGlyGlnTrpGlu 162
Db 883 AGCCATTCTGTGCGCTTT-----GCAAACTGGAGACGGCAA 918
Qy 163 AlaGluAspLeuHis-----LeuAspAspGlu 171
Db 919 ATCTCTGAGCTTCATCCAAATTTGCAACAAATCTATTTTAAAGCAAGAAGTTGATTATATG 978
Qy 172 GluSerSerLysArgProLeuGlyLeuLeuAlaArgGlnAlaGluAsnHisTyrAspGln 191
Db 979 ACTCAGGCTAGGGGTACAGAGATCTCTCTGGCAGAAGACAATGAGTCAGCTACACGAGA 1038
Qy 192 AspLeuAspGluLeuGlnLeuGluMetGluAspSerLysProHisPro-----SerVal 209
Db 1039 GGATTGTGACATGACGTACACTGAGTTTGACTATGACTTATGCAATGAAGTGGTGCAGCTG 1098
Qy 210 GlnCysSerProThrProGlyProPheLysProCysGluTyrLeuPheGluSerTrpGly 229
Db 1099 ACCTGCTCCCTAAGCCAGATGCATTCACCCATGTGAAGATATCATGGGTACACATC 1158
Qy 230 IleArgLeuAlaValTrpAlaIleValLeuLeuSerValLeuCysAsnGlyLeuValLeu 249
Db 1159 CTCAGAGCTCGATATGTTTATCAGCATCTCTGGCCATCACTGGGAAC---ATCATAGTG 1215
Qy 250 LeuThrValPheAlaGlyGlyProAlaProLeuProValLysPheValValGlyAla 269
Db 1216 CTAGTGATCCCTAACTACACGCAATATATAAATCAACAGTCCCGAGTTCTTATGTGCAAC 1275
Qy 270 IleAlaGlyAlaAsnThrLeuThrGlyIleSerCysGlyLeuLeuAlaSerValAspAla 289
Db 1276 CTGGCTTTGTGATCTCTGATTTGAATCTACCTGCTGCTCATGTGCATCAGTTGATATC 1335
Qy 280 LeuThrPheGlyGlnPheSerGluTyrGlyAlaArgTrpGluThrGlyLeuGlyCysArg 309
Db 1336 CATACCAAGAGCCAAATATACAACTATGCAATGTAGCTGGCAAACTGGGGCAGGCTGTGAT 1395
Qy 310 AlaThrGlyPheLeuAlaValLeuGlySerGluAlaSerValLeuLeuLeuThrLeuAla 329
Db 1396 GCTGCTGCTTTTTCACCTGTCTTTGCCAGTGAGTGTGCTACACTCTGACAGCTATC 1455
```

```
Qy 330 AlaValGln-----CysSerValSerValSerCysValArgAlaTyr 343
Db 1456 ACCTTGGAAAGATGGCATACCATCGCATGCCATGCCATGGACTGGCTGC----- 1503
Qy 344 GlyLysSerProSerLeuGlySerValArgAlaGlyValLeuGlyCysLeuAlaLeuAla 363
Db 1504 ---AAGGTGACGCTCCGCCATGCTGCCAGTGTCTATGGTGTGGGTGGATTT---TTTGCT 1557
Qy 364 GlyLeuAlaAlaAlaLeuProLeuAlaSerValGlyGluTyrGlyAlaSerProLeuCys 383
Db 1558 TTTGCAGCTGCCCTCTTTCCCATCTTTGGCATCAGCAGCTACATGAGGTGAGCATCTGC 1617
Qy 384 LeuProTyrAlaProGluGlyGlnProAlaAlaLeuGlyPheThrValAlaLeuVal 403
Db 1618 CTGCCCATG-----GATATTGACAGCCCTTTGTCACAGCTGTATGTCTGCCCTCCTT 1671
Qy 404 MetMetAsnSerPheCysPheLeuValValAlaGlyAlaTyrIleLysLeuTyrCysAsp 423
Db 1672 GTGCTCAATGCTCGGCCCTTTGTGTCATCTGTGGCTGTATATCCATCTACCTCACA 1731
Qy 424 LeuProArgGlyAspPhe---GluAlaValTrpAspCysAlaMetValArgHisValAla 442
Db 1732 GTGGCGGAACCCACATCGTGTCTCTCTAGTGACACAGCATGCCCAAGCGCATGCC 1791
Qy 443 TrpLeuIlePheAlaAspGlyLeuLeuTyrCysProValAlaPheLeuSerPheAlaSer 462
Db 1792 ATGCTCATCTTCACTGACTTCTCTGATGACGCCCATTTCTTTCTTTGCGCATTTCTGCC 1851
Qy 463 MetLeuGlyLeuPheProValThrProGluAlaValLysSerValLeuValValLeu 482
Db 1852 TCCCTCAAGGTGCCCCATCATCTGTGTCCAAAGAAAGATTTGCTGGTCTTGTTTCCAC 1911
Qy 483 ProLeuProAlaCysLeuAsnProLeuLeuLeuTyrLeuLeuPheAsnProHisPheArgAsp 502
Db 1912 CCCATCACTCTGTGCCAACCCCTTCTCTATGCCATCTTTACCAAAAACCTTTTCGCGAGA 1971
Qy 503 AspLeuArgArgLeuArgProArgAlaGlyAspSerGlyProLeuAlaTyrAlaAlaAla 522
Db 1972 GATTCTTCACTTCTGCTGAGCAAGTGGC-----TGCTATGAAATGCAA 2016
Qy 523 GlyGluLeuGlyLysSerSerCysAspSerThr 533
Db 2017 GCCCAATTTATAGGACAGAAAACCTTCATCCACT 2049
RESULT 15
; Sequence 57, Application US/07741453A
; Patent No. 6228597
; GENERAL INFORMATION:
; APPLICANT: PARMENTIER, MARC
; APPLICANT: LIBERT, FREDERIC
; APPLICANT: DUMONT, JACQUES
; APPLICANT: VASSART, GILBERT
; TITLE OF INVENTION: POLYPEPTIDES HAVING THYROTROPIN-RECEPTOR
; TITLE OF INVENTION: ACTIVITY, NUCLEIC ACID SEQUENCES CODING FOR SUCH RECEPTORS
; TITLE OF INVENTION: AND POLYPEPTIDES, AND APPLICATIONS OF THESE POLYPEPTIDES
; NUMBER OF SEQUENCES: 62
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: CUSHMAN DABY & CUSHMAN
; STREET: 1615 L STREET, N.W.
; CITY: WASHINGTON, D.C.
; COUNTRY: U.S.A.
; ZIP: 20036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07741,453A
; FILING DATE: 19911015
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
```

```

; NAME: KOKULIS, PAUL N.
; REGISTRATION NUMBER: 16773
; REFERENCE/DOCKET NUMBER: 91913/1107/US/ST
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 861-3000
; TELEFAX: (202) 822-0944
; TELEX: 6714627 CUSH
; INFORMATION FOR SEQ ID NO: 57:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4417 base pairs
; TYPE: NUCLEIC ACID
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; US-07-741-453A-57

Alignment Scores:
Pred. No.: 6,21e-29 Length: 4417
Score: 414.50 Matches: 141
Percent Similarity: 40.77% Conservative: 102
Best Local Similarity: 23.66% Mismatches: 258
Query Match: 12.53% Indels: 95
DB: 3 Gaps: 12

US-10-664-667-5 (1-633) x US-07-741-453A-57 (1-4417)

Qy 33 SerLeuProAlaHisProAlaSerLeuAlaAlaAlaSerAsnThrThrAlaSer 52
Db |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
558 TCCATCCCTCGCAAT-----GCTTCCAGGGGCTGTGCAATGAACCCCTGACA 605
Qy 53 GlyLysLeuGlu***AspThrPheSerGlnLeu-----SerSer 65
Db |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
606 CTGAACATATACAAATGGCTTACTTCAATCCAGGACATGCTTCAATGGGACAAA 665
Qy 66 LeuGlnAlaLeuAspLeuSerTrpAsnAla----IleArgSerIleHisProGluAlaPhe 84
Db |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
666 CTGATGCTGTTTACCTGAACAAGAATAAATACCTGTGAGTATCGAATAAGATGCAATT 725
Qy 85 SerThrLeuHisSer----LeuValLysLeuAspLeuThrAspAsnGlnLeuThrLieu 103
Db |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
726 GGAGGAGTGATACAGTGGACCAACCTTGTGATGCTCTTACACAGGTGTACTGCGCTG 785
Qy 104 ProLeuAlaGlyLeuGlyGlyLeuMetHisLysLysLeuLysGlyAsnLeuAlaLeuSer 123
Db |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
786 CCATCCAAAGCCCTGGAGCATCTAAAGAGCTGATAGCAAGAACACATTGACCTCTAAG 845
Qy 124 GlnAlaPheSerLysAspSerPheProLysLeuArgIleLeuGluValProTyrAlaTyr 143
Db |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
846 AAACCTCCACTTTCCTTTGAGTTTCTTACCTTACACGGGCTGACCTTCTTATCCAAGC 905
Qy 144 GlnCysCysProTyr-----GlyMetCysAlaSerPhe----- 154
Db |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
906 CACTGTGTGCTTTTAAAGAAATCAGAAGAAATCAGAGGAATCCTTGAATGCTTAAATGTG 965
Qy 154 ----- 154
Db
Qy 966 AATCAAGACAGTATTCGAGGCTCGCCAGAGAAAATCTGTGAATACTTTGAATGGCCCC 1025
Qy 155 -----PheLysAlaSerGly 159
Db |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1026 TTTGACCAGGAATATGAAGAGTATCTGGGTGACAGCCATGCTGGGTACAAGGCAACTCT 1085
Qy 160 GlnTrp-----GluAlaGluAsp 165
Db |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1086 CAGTTCAGGATACCGATAGCAATTCATTTATGCTTCTTTCGAAGAACCAAGAGAT 1145
Qy 166 LeuHisLeuAspAspGluGluSerSerLysArgProLeuGlyLeuLeuAlaArgGlnAla 185
Db |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1146 GAGATCTTCGGTTTGGGAGGAGCTTAAACCCAGGAAGACCCCTCCAGGCCCTT 1205
Qy 186 GluAsnHisTyrAspGlnAspLeuAspGluLeuGlnLeuMetGluAspSerLysPro 205
Db |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1206 GATAGCCATTATGACTACACTGTG-----TGTTGGTGGC 1238

```

```

Qy 206 HisProSerValGlnCysSerProThrProGlyProPheLysProCysGluTyrLeuPhe 225
Db |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1239 AATGAAGACATGGTGTACTTCAAGTCAGATCAGATTCAACCCCTTGAAGACATATG 1298
Qy 226 GluSerTrpGlyIleArgLeuAlaValTrpAlaIleValLeuLeuSerValLeuCysAsn 245
Db |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1299 GGCTACAAGTTCCTGAGGATGTGGTGTGGTTGTAGTCTGCTGGCTCTCTCTGGGCAT 1358
Qy 246 GlyLeuValLeuLeuThrPheAlaGlyGlyProAlaProLeuProProValLysPhe 265
Db |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1359 GTCTTCTCTGATCGCTCTCTTACCAGTCACTACAAATGACTGTCCCA---CGCTTT 1415
Qy 266 ValValGlyAlaIleAlaGlyAlaAsnThrLeuThrGlyIleSerCysGlyLeuLeuAla 285
Db |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1416 CTATGTGCAACTTGGCTTTGCAGATTTCTGCATGGGAGTATCTGCCTCTCATCGCC 1475
Qy 286 SerValAspAlaLeuThrPheGlyGlnPheSerGluTyrGlyAlaArgTrpGluThrGly 305
Db |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1476 TCCGTAGACCTCTACACTCATTCTGAGTACTACAAACCATGCCATCGACTGGCAGCAGGC 1535
Qy 306 LeuGlyCysArgAlaThrGlyPheLeuAlaValLeuGlySerGluAlaSerValLeuLeu 325
Db |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1536 CTTGGGTGTAACACAGCTGTTCTTCTCAGTGTCTTTCAGTGAATATCAGTGTATACA 1595
Qy 326 LeuThrLeuAlaValGlnCysSerValSerValSerCysValArgAlaTyrGlyLys 345
Db |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1596 CTGACAGTCAATCCCTGGAGCGCTGGTATGCCATTACCTTTCGCATCGCGCTGGACAGG 1655
Qy 346 SerProSerLeuGlySerValArgAlaGlyValLeuGlyCysLeuAlaLeuAlaGlyLeu 365
Db |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1656 AAGATCGCTTCAGGCATGCATATGCCATCATGTTGGGGCTGGGTTGCTGCTTCTGT 1715
Qy 366 AlaAlaAlaLeuProLeuAlaSerValGlyGlyTyrGlyAlaSerProLeuCysLeuPro 385
Db |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1716 CTGCGCTGCTCCTCTGTTGGGAATAAGCAGCATATGCCAAGGTGACATCTGCTGCC 1775
Qy 386 TyrAlaProGluGlyGlnProAlaAlaLeuGlyPheThrValAlaLeuValMetMet 405
Db |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1776 ATG-----GACACTGAGACACCTTGGCCCTGGCATATATATCTTGTGTTGCTC 1829
Qy 406 AsnSerPheCysPheLeuValValAlaGlyAlaTyrIleLysLeuTyrCysAspLeuPro 425
Db |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1830 AACATAGTTCCTTTATCATGTTCTGCTCTGTTATGTGAAGATCTACATCACAGTCCGA 1889
Qy 426 ArgGlyAspPheGlu---AlaValTrpAspCysAlaMetValArgHisValAlaTrpLeu 444
Db |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1890 AATCCCGCAGTACAAACCCGGGGACAAAGACACCAAAATTCGCAAAAGGATGGCTGATTG 1949
Qy 445 IlePheAlaAspGlyLeuLeuTyrCysProValAlaPheLeuSerPheAlaSerMetLeu 464
Db |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1950 ATCTTCACTGATTCATGTCATGTCATGCGCCCAATCTCATCTACGCTCTGTGACGACTTATG 2009
Qy 465 GlyLeuPheProValThrProGluAlaValLysSerValLeuLeuValValLeuProLeu 484
Db |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
2010 AACAGCCTCTCATCTACTTACCAACTCCAAATCTTGCTGGTCTCTCTTCTTATCACTT 2069
Qy 485 ProAlaCysLeuAsnProLeuLeuTyrLeuLeuPheAsnProHisPheArgAspLeu 504
Db |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
2070 AACTCTGTGCCAATCCATTTCTATGCTATTTTCCAGAAAGCCTTCCAGAGGAGTATA 2129
Qy 505 ArgArgLeuArgProArgAlaGlyAspSerGlyProLeuAlaTyrAlaAlaGly--- 523
Db |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
2130 TTTATCTGCTCAGCAAGTTTGGCATCTGTAAAGCCGAGCTCAGGCATCCCGGGGCAG 2189
Qy 524 -----GluLeuGlyLysSerSerCysAspSerThr 533
Db |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
2190 AGGGTTTCTCCAAAGATAGTGTGTTATTCAGATCCAAAGGTTTACCCGGGACATGAGG 2249
Qy 534 GlnAlaLeuValAlaPheSerAspValAspLeuLeuGluAlaSer 549
Db |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
2250 CAAAGTCTCCCAACATGAGGATGAGTATGAATGCTTGAATACTCC 2297

```

Search completed: April 14, 2005, 21:52:31
Job time : 256 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model
Run on: April 14, 2005, 19:13:44 ; Search time 637 Seconds
(without alignments)
6028.616 Million cell updates/sec

Title: US-10-664-667-5
Perfect score: 3307
Sequence: 1 NTHYRESWACRYRSGIPG.....GGLSGGGGFGPQLAFASHV 633

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 5622541 seqs, 303335566 residues

Total number of hits satisfying chosen parameters: 11245082

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+p2n.model -DEV=xlh
-Q=/cgn2_1/USPTO_spool/US10664667/runat_13042005_184540_7329/app_query.fasta_1.775
-DB=Published_Applications_NA -QFM=fastap -SUFFIX=p2n.rnpb -MINMATCH=0.1
-LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blomum62
-TRANS=human40.cdi -LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100
-THR MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0
-MAXLEN=2000000000 -USER=US10664667 @CEN 1.1 354 @runat_13042005_184540_7329
-NCPU=6 -ICPU=3 -NO MAP -LARGEQUERY -NEG SCORES=0 -WAIT DSPBLOCK=100
-LONGLOG -DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5
-FGAPOP=6 -FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Published Applications NA.*

- 1: /cgn2_6/ptodata/1/pubna/US07_PUBCOMB.seq.*
- 2: /cgn2_6/ptodata/1/pubna/PCT_NEW_PUB.seq.*
- 3: /cgn2_6/ptodata/1/pubna/US06_NEW_PUB.seq.*
- 4: /cgn2_6/ptodata/1/pubna/US06_PUBCOMB.seq.*
- 5: /cgn2_6/ptodata/1/pubna/US07_NEW_PUB.seq.*
- 6: /cgn2_6/ptodata/1/pubna/PCTUS_PUBCOMB.seq.*
- 7: /cgn2_6/ptodata/1/pubna/US08_NEW_PUB.seq.*
- 8: /cgn2_6/ptodata/1/pubna/US08_PUBCOMB.seq.*
- 9: /cgn2_6/ptodata/1/pubna/US09A_PUBCOMB.seq.*
- 10: /cgn2_6/ptodata/1/pubna/US09B_PUBCOMB.seq.*
- 11: /cgn2_6/ptodata/1/pubna/US09C_PUBCOMB.seq.*
- 12: /cgn2_6/ptodata/1/pubna/US09_NEW_PUB.seq.*
- 13: /cgn2_6/ptodata/1/pubna/US10A_PUBCOMB.seq.*
- 14: /cgn2_6/ptodata/1/pubna/US10B_PUBCOMB.seq.*
- 15: /cgn2_6/ptodata/1/pubna/US10C_PUBCOMB.seq.*
- 16: /cgn2_6/ptodata/1/pubna/US10D_PUBCOMB.seq.*
- 17: /cgn2_6/ptodata/1/pubna/US10E_PUBCOMB.seq.*
- 18: /cgn2_6/ptodata/1/pubna/US10F_PUBCOMB.seq.*
- 19: /cgn2_6/ptodata/1/pubna/US10_NEW_PUB.seq.*
- 20: /cgn2_6/ptodata/1/pubna/US11_NEW_PUB.seq.*
- 21: /cgn2_6/ptodata/1/pubna/US60_NEW_PUB.seq.*
- 22: /cgn2_6/ptodata/1/pubna/US60_PUBCOMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query Match	Score	Length	DB ID	Description
1	99.9	3305	1899	10	US-09-851-595-6
2	99.9	3305	1899	17	US-10-664-667-6
3	99.9	3305	2486	10	US-09-851-595-4
4	99.9	3305	2486	17	US-10-664-667-4
5	92.1	3045.5	2208	10	US-09-851-595-9
6	92.1	3045.5	2208	17	US-10-664-667-9
7	92.1	3045.5	2711	10	US-09-851-595-7
8	92.1	3045.5	2711	17	US-10-664-667-7
9	92.1	3045.5	2901	10	US-09-851-595-12
10	92.1	3045.5	2901	17	US-10-664-667-12
11	92.1	3045.5	3119	15	US-10-325-567A-580
12	92.1	3045.5	3325	17	US-10-331-496A-52
13	92.1	3045.5	3429	18	US-10-398-036-26
14	92.1	3045.5	3438	14	US-10-176-847-89
15	92.1	3045.5	3443	17	US-10-331-496A-94
16	92.1	3045.5	3492	10	US-09-851-595-10
17	92.1	3045.5	3492	17	US-10-664-667-10
18	92.1	3045.5	3492	18	US-10-737-450-31
19	92.0	3041.5	3273	17	US-10-295-027-1119
20	92.0	3041.5	3273	18	US-10-783-528-43
21	91.9	3040.5	2487	14	US-10-270-336-1
22	91.9	3040.5	2988	17	US-10-302-172-723
23	91.9	3038.5	3349	14	US-10-370-336-4
24	91.2	3016.5	3381	10	US-09-970-944-5
25	79.8	2638	2901	10	US-09-851-595-3
26	79.8	2638	2901	17	US-10-664-667-3
27	79.8	2638	3637	17	US-09-851-595-1
28	79.8	2638	3637	17	US-10-664-667-1
29	66.3	2193	119596	14	US-10-270-336-3
30	1275	38.6	2651	17	US-10-295-027-483
31	1275	38.6	2651	17	US-10-295-027-1113
32	1275	38.6	2651	17	US-10-173-999-27
33	1275	38.6	2724	15	US-10-251-385-263
34	1275	38.6	2724	15	US-10-225-567A-421
35	1275	38.6	2724	15	US-10-174-456-4
36	1275	38.6	2724	18	US-10-751-736-21
37	1275	38.6	2880	17	US-10-295-027-1114
38	1275	38.6	2880	17	US-10-482-029-157
39	1275	38.6	2880	19	US-10-651-237-48
40	1275	38.6	2880	19	US-10-782-413-48
41	1275	38.6	3032	17	US-10-295-027-848
42	1275	38.6	3032	17	US-10-395-027-945
43	1273	38.5	2082	19	US-10-851-470-3
44	1269	38.4	2724	15	US-10-251-385-277
45	1242.5	37.6	1173	17	US-10-343-650A-65

ALIGNMENTS

RESULT 1
US-09-851-595-6
; Sequence 6, Application US/09851595
; Publication No. US20030166047A1
; GENERAL INFORMATION:
; APPLICANT: Gl, Wei
; TITLE OF INVENTION: NOVEL G-PROTEIN COUPLED RECEPTORS AND USES THEREFOR
; FILE REFERENCE: MNI-080CP
; CURRENT APPLICATION NUMBER: US/09/851,595
; CURRENT FILING DATE: 2000-05-08
; PRIOR APPLICATION NUMBER: 09/556,588
; PRIOR FILING DATE: 2000-05-08
; PRIOR APPLICATION NUMBER: 60/132,896
; PRIOR FILING DATE: 1999-05-06
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 6
; LENGTH: 1899
; TYPE: DNA
; ORGANISM: Homo sapiens

```
;
; NAME/KEY: CDS
; LOCATION: (1) .. (1899)
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (171)
; OTHER INFORMATION: n = any nucleotide
US-09-851-595-6

Alignment Scores:
Pred. No.: 0 Length: 1899
Score: 3305.00 Matches: 632
Percent Similarity: 99.84% Conservative: 0
Best Local Similarity: 99.84% Mismatches: 1
Query Match: 99.94% Indels: 0
DB: 10 Gaps: 0

US-10-664-667-5 (1-633) x US-09-851-595-6 (1-1899)

Qy 1 AenThrThrHisTyrArgGluSerTrpTyrAlaCysArgTyrArgSerGlyLeuProGly 20
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy 1 AATAGCACTACTATAGGAAAGCTGGTACGCTGCAGGTACCGGTCCGGAATTCCTCGGG 60
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy 21 SerThrHisAlaSerValGluArgSerGlnGlyLeuSerLeuProAlaHisProAlaSer 40
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy 61 TCGACCCACGCGTCCGTGAGCGGAGCCAGGGTCTGAGCCTCGCGGTCTATCCAGCCTCT 120
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy 41 LeuAlaAlaLeuAlaAlaSerAsnThrThrAlaSerGlyLysLeuGlu***AspThrPhe 60
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy 121 CTGTGCGCCTAGCGGCTCCACACACACACACACACACACACACACACACACACCTTC 180
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy 61 SerGlnLeuSerSerLeuGlnAlaLeuAspLeuSerTrpAsnAlaIleArgSerIleHis 80
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy 181 AGCAGCTGAGCTCCCTGCAAGCCCTGGATCTTACGTGGAACGCCATCCGGTCCATCCAC 240
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy 81 ProGluAlaPheSerThrLeuHisSerLeuValLysLeuAspLeuThrAspAsnGlnLeu 100
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy 241 CCTGAGGCTTCTCCACCTGCACTCCCTGCTGCTGAGTGGACCTGACACACACAGCTG 300
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy 101 ThrThrLeuProLeuAlaGlyLeuGlyLeuMethHisLeuLysLeuLysGlyAsnLeu 120
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy 301 ACCACACTGCCCCTGGCTGGACTTTGGGGCTTGATGCATCTGAAGCTCAAGAGGAACTT 360
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy 121 AlaLeuSerGlnAlaPheSerLysAspSerPheProLysLeuArgIleLeuGluValPro 140
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy 361 GCTCTCTCCAGGCTTCTCCAGAGACAGTTTCCCAAAACTGAGGATCCTGGAGGTGCTT 420
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy 141 TyrAlaTyrGlnCysCysProTyrGlyMetCysAlaSerPhePheLysAlaSerGlyGln 160
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy 421 TATGCTACCACTGCTGCTCCCTATGGGATGTGTGCCAGCTTCTTCAAGGCTCTTGGGCAG 480
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy 161 TrpGluAlaGluAspLeuHisLeuAspAspGluGluSerSerLysArgProLeuGlyLeu 180
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy 481 TGGGAGGCTGAAGACCTTCACTTGGATGATGAGGAGTCTTCAAAAAGGCCCTGGGCCTC 540
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy 181 LeuAlaArgGlnAlaGluAsnHisTyrAspGlnAspLeuAspGluLeuGlnLeuMet 200
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy 541 CTTGCCAGACAAAGCAGAGAACCTATATGACCAGGACCTGGATGAGCTCCAGCTGGAGATG 600
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy 201 GluAspSerLysProHisProSerValGlnCysSerProThrProGlyProPheLysPro 220
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy 601 GAGGACTCAAAGCCACACCCAGTGTCCAGTGTAGCCCTTACTCCAGGCCCTTCAAGCCC 660
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy 221 CysGluTyrLeuPheGluSerTrpGlyLeuArgLeuAlaValTrpAlaIleValLeuLeu 240
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy 661 TGTGAGTACCTCTTTGAAAGCTGGGGCATCCGCTGCGGTGGGCGCATCGTGTGCTC 720
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy 241 SerValLeuCysAsnGlyLeuValLeuLeuThrValPheAlaGlyGlyProAlaProLeu 260
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy 721 TCCGTGCTCTGCAATAGACTGGTGTGCTGACCGTGTGCTGGCGGCTGCCCCCTG 780
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy 261 ProProValLysPheValValGlyAlaIleAlaGlyAlaAsnThrLeuThrGlyLeuSer 280
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db 781 CCCCCGCTCAAGTTTGTGGTAGGTGCGATTGACGGCGCAACACACTTGACTGGCATTTCC 840
Qy 281 CysGlyLeuLeuAlaSerValAspAlaLeuThrPheGlyGlnPheSerGluTyrGlyAla 300
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 841 TGTGGCCTTCTAGCCTCAGTCGATCGCCTGACCTTTGTGTGTCAGTTCTCTGAGTAGCGAGCC 900
Qy 301 ArgTrpGluThrGlyLeuGlyCysArgAlaThrGlyPheLeuAlaValLeuGlySerGlu 320
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 901 CGCTGGGAGACGGGGCTAGGCTGCCGGGCCACTGGCTTCTGGCAGTACTTGGGTGGAG 960
Qy 321 AlaSerValLeuLeuLeuThrLeuAlaAlaValGlnCysSerValSerValSerCysVal 340
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 961 GCATCGTGTGCTGCTCACTCTGGCGCAGTCAGTCAGTCGCTCCGTCCTCTGCTGTC 1020
Qy 341 ArgAlaTyrGlyLysSerProSerLeuGlySerValArgAlaGlyValLeuGlyCysLeu 360
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1021 CGGGCCTATGGGAAGTCCCTCTCCCTGGGAGCGCTTCGAGCAGGGGTCTTAGGTGCGTG 1080
Qy 361 AlaLeuAlaGlyLeuAlaAlaAlaLeuProLeuAlaSerValGlyGluTyrGlyAlaSer 380
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1081 GCATCGAGGGCTGGCGCGCACTGCCCTCGGCTCAGTGGAGAAATACGGGGCCTTCC 1140
Qy 381 ProLeuCysLeuProTyrAlaProProGluGlyGlnProAlaAlaLeuGlyPheThrVal 400
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1141 CCACCTCTGCTGCCCTACGCCACCTGAGGGTCAGCCAGCAGCCCTGGGCTTCCACGCTG 1200
Qy 401 AlaLeuValMetMetAsnSerPheCysPheLeuValValAlaGlyAlaTyrIleLysLeu 420
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1201 GCCCTGCTGATGATGAATCTCTTCTGTCTGCTGCTGCGCGGTGCTACATCAAACTG 1260
Qy 421 TyrCysAspLeuProArgGlyAspPheGluAlaValTrpAspCysAlaMetValArgHis 440
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1261 TACTGTGACCTGCCCGGGGGAGCTTTGAGCGCGTGGGAGCTCGCCCATGGTGAGGCAC 1320
Qy 441 ValAlaTrpLeuIlePheAlaAspGlyLeuLeuTyrCysProValAlaPheLeuSerPhe 460
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1321 GTGGCCTGGCTCATCTTCGCAGACGGGCTCCTCTACTGTCCCGTGGCTTCTCTCAGCTTC 1380
Qy 461 AlaSerMetLeuGlyLeuPheProValThrProGluAlaValLysSerValLeuVal 480
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1381 GCCTCCATGTGGGCTCTTCCCTGTGCAGCCGAGCGCTCAAGTCTGTCTCTCTGCTG 1440
Qy 481 ValLeuProLeuProAlaCysLeuAsnProLeuTyrLeuLeuPheAsnProHisPhe 500
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1441 GTGCTGCCCTGCTGCTGCTGCCCTCAACCCACTGCTGTACTGTCTTCAACCCCACTTC 1500
Qy 501 ArgAspAspLeuArgArgLeuArgProArgAlaGlyAspSerGlyProLeuAlaTyrAla 520
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1501 CGGGATGACCTTCGGGGCTTCGGCCCCCGCAGGGGAGCTCAGGGCCCCCTAGCCTATGCT 1560
Qy 521 AlaAlaGlyGluLeuGluLysSerSerCysAspSerThrGlnAlaLeuValAlaPheSer 540
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1561 GCGGCGGGAGCTGGAGAAAGACTCCTGTGTATTTCTACCCAGGCTTGGTGGCTTCTCT 1620
Qy 541 AspValAspLeuIleLeuGluAlaSerGluAlaGlyArgProProGlyLeuGluThrTyr 560
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1621 GATGTGATCTCATTTCTGAAGCTTTCTGAAGCTGGCGGCCCTCTGGGCTGGAGACTAT 1680
Qy 561 GlyPheProSerValThrLeuIleSerCysGlnGlnProGlyAlaProArgLeuGluGly 580
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1681 GGCTTCCCTCTCAGTGACCTCATCTCTGTTCAGCAGCAGCGGGCCCCCAGGCTGGAGGGC 1740
Qy 581 SerHisCysValGluProGluGlyAsnHisPheGlyAsnProGlnProSerMetAspGly 600
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1741 AGCCATTTGTAGAGCCAGAGGGGAACCACTTTGGGAACCCCAACCTCTCATGGATGGA 1800
Qy 601 GluLeuLeuLeuArgAlaGluGlySerThrProAlaGlyGlyLeuSerGlyGlyGly 620
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1801 GAATCTGCTGAGGGCAGAGGATCTTACCCAGAGCTGGAGCTTGTGAGGGGTGGC 1860
Qy 621 GlyPheGlnProSerGlyLeuAlaPheAlaSerHisVal 633
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1861 GGCTTTCAGGCCCTCTGGCTTGGCTTTGCTTTCACAGCTG 1899
```

RESULT 2

US-10-664-667-6
 ; Sequence 6, Application US/10664667
 ; Publication No. US20040058377A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Gu, Wei
 ; TITLE OF INVENTION: NOVEL G-PROTEIN COUPLED RECEPTORS AND USES THEREFOR
 ; FILE REFERENCE: MNI-080CP
 ; CURRENT APPLICATION NUMBER: US/10/664,667
 ; CURRENT FILING DATE: 2003-09-18
 ; PRIOR APPLICATION NUMBER: US/09/851,595
 ; PRIOR FILING DATE: 2000-05-08
 ; PRIOR APPLICATION NUMBER: 09/556,588
 ; PRIOR FILING DATE: 2000-05-08
 ; PRIOR APPLICATION NUMBER: 60/132,896
 ; PRIOR FILING DATE: 1999-05-06
 ; NUMBER OF SEQ ID NOS: 12
 ; SOFTWARE: PatentIn ver. 2.0
 ; SEQ ID NO 6
 ; LENGTH: 1899
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; FEATURE:
 ; NAME/KEY: CDS
 ; LOCATION: (1)..(1899)
 ; FEATURE:
 ; NAME/KEY: misc feature
 ; LOCATION: (171)
 ; OTHER INFORMATION: n = any nucleotide
 US-10-664-667-6

Alignment Scores:

Pred. No.: 0 Length: 1899
 Score: 3305.00 Matches: 632
 Percent Similarity: 99.84% Conservatives: 0
 Best Local Similarity: 99.84% Mismatches: 1
 Query Match: 99.94% Indels: 0
 DB: 17 Gaps: 0

US-10-664-667-5 (1-633) x US-10-664-667-6 (1-1899)

Qy	1	AsnThrHisTyrArgGluSerTrpTyrAlaCysArgTyrArgSerGlyIleProGly	20
Db	1	AATACGACTCATATAGGAAAGCTGGTACGCTCGAGGTACCGTCCGGAAATCCCGG	60
Qy	21	SerThrHisAlaSerValArgSerGlnGlyLeuSerLeuProAlaHisProAlaSer	40
Db	61	TCGACCCACGGCTCGTGGAGCGGAGCCAGGCTCTGAGCCTGCGGCTCATCCAGCCTCT	120
Qy	41	LeuAlaAlaLeuAlaAlaSerAsnThrThrAlaSerGlyLysLeuGlu***AspThrPhe	60
Db	121	CTTGCTGCCCTAGCGGCTCCACACACACCGCATCTGGGAAATGGAGCTNGACACTTC	180
Qy	61	SerGlnLeuSerSerLeuGlnAlaLeuAspLeuSerTrpAsnAlaIleArgSerIleHis	80
Db	181	AGCCAGCTGAGTCCCTCTCAAGCCCTGGATCTTAGCTGGAACGCCATCCGGTCCATCCAC	240
Qy	81	ProGluAlaPheSerThrLeuHisSerLeuValLysLeuAspLeuThrAspAsnGlnLeu	100
Db	241	CCTGAGGCTTCTCCACCCCTGACCTCCCTGGTCAAGCTGGAGCTTGACAGCAACACAGCTG	300
Qy	101	ThrThrLeuProLeuAlaGlyLeuGlyLeuMetHisLeuLysLeuLysGlyAsnLeu	120
Db	301	ACCACACTGCCCTGGCTGGACTTGGGGCTTGATGCACTTGAAGCTCAAAGGNAACCTT	360
Qy	121	AlaLeuSerGlnAlaPheSerLysAspSerPheProLysLeuArgIleLeuGluValPro	140
Db	361	GCTCTCTCCAGGCCCTTCTCAAGACACAGTTCCTCCCAAACTGAGGATCCTGGAGGTGCT	420
Qy	141	TyrAlaTyrGlnCysCysProTyrGlyMetCysAlaSerPhePheLysAlaSerGlyGln	160
Db	421	TATGCTTACCAGTGTGTCCCTATGGGATGTGTGCCAGCTTCTTCAAGGCCCTCTGGGAG	480

Qy	161	TrpGluAlaGluAspLeuHisLeuAspAspGluSerSerLysArgProLeuGlyLeu	180
Db	481	TGGGAGGCTGAAGACCTTACCTTGATGATGAGGAGTCTTCAAAAAGCCCTCGGCTC	540
Qy	181	LeuAlaArgGlnAlaGluAsnHisTyrAspGlnAspLeuAspGluLeuGluLeuMet	200
Db	541	CTTCCAGACAGAGCAGAGAACCTATGACAGACCTGGATGAGCTCCAGCTGGAGATG	600
Qy	201	GluAspSerLysProHisProSerValGlnCysSerProThrProGlyProPheLysPro	220
Db	601	GAGGACTCAAAGCCACACCCAGTGTCCAGTGTAGCCCTACTCCAGGCCCTTCAAGCCC	660
Qy	221	CysGluTyrLeuPheGluSerTrpGlyIleArgLeuAlaValTrpAlaIleValLeuLeu	240
Db	661	TGTGAGTACCTCTTTGAAAGCTGGGCATCCGCTGCGCGTGTGGCCATCGTGTGTGCT	720
Qy	241	SerValLeuCysAsnGlyLeuValLeuLeuThrValPheAlaGlyGlyProAlaProLeu	260
Db	721	TCCGTGCTCTGCAATGGACTGGTGTCTGCTGACCTGTTCTGCTGGCGGCCCTGCCCCCTG	780
Qy	261	ProProValLysPheValValGlyAlaIleAlaGlyAlaAsnThrLeuThrGlyIleSer	280
Db	781	CCCCGGTCAAGTTTGTGTAGTCCGATTGACAGGCCCAACACCTTGACTGGCATTTCC	840
Qy	281	CysGlyLeuLeuAlaSerValAspAlaLeuThrPheGlyGlnPheSerGluTyrGlyAla	300
Db	841	TGTGGCTTCTAGCTCAGTCCGCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	900
Qy	301	ArgTrpGluThrGlyLeuGlyCysArgAlaThrGlyPheLeuAlaValLeuGlySerGlu	320
Db	901	CGCTGGGAGACGGGCTAGGCTGCGGGCCACTGCTTCTGGCAGTAGTCTTGGGTCGAG	960
Qy	321	AlaSerValLeuLeuLeuThrLeuAlaAlaValGlnCysSerValSerValSerCysVal	340
Db	961	GCATCGGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	1020
Qy	341	ArgAlaTyrGlyLysSerProSerLeuGlySerValArgAlaGlyValLeuGlyCysLeu	360
Db	1021	CGGCTTATGGAAAGTCCCTCTGCGGAGCGTTCGAGCAGGGCTCTAGGCTGCTGCTG	1080
Qy	361	AlaLeuAlaGlyLeuAlaAlaAlaLeuProLeuAlaSerValGlyGlyAlaSer	380
Db	1081	GCATCGGAGGCTGGCGCGCCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	1140
Qy	381	ProLeuCysLeuProTyrAlaProProGluGlnProAlaAlaLeuGlyPheThrVal	400
Db	1141	CCACTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	1200
Qy	401	AlaLeuValMetMetAsnSerPheCysPheLeuValValAlaGlyAlaTyrIleLysLeu	420
Db	1201	GCCCTGGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG	1260
Qy	421	TyrCysAspLeuProArgGlyAspPheGluAlaValTrpAspCysAlaMetValArgHis	440
Db	1261	TACTGTGACCTGCGCGGGGAGCTTTGAGGCCGTGTGGGAGCTGCGGCATGTTGAGGAC	1320
Qy	441	ValAlaTrpLeuIlePheAlaAspGlyLeuLeuTyrCysProValAlaPheLeuSerPhe	460
Db	1321	GTGCGCTGGCTGATCTTCTGCGAGAGGGCTCTCTACTGCTCCGCGGCTTCTTCTAGCTTC	1380
Qy	461	AlaSerMetLeuGlyLeuPheProValThrProGluAlaValLysSerValLeuLeuVal	480
Db	1381	GCCTCCATGCTGGGCTCTTCTGCTGTCAGCCCGGAGGCCGTCAAGTGTCTGCTGCTG	1440
Qy	481	ValLeuProLeuProAlaCysLeuAsnProLeuLeuTyrLeuLeuPheAsnProHisPhe	500
Db	1441	GTGCTGCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	1500
Qy	501	ArgAspLeuArgArgLeuArgProArgAlaGlyAspSerGlyProLeuAlaTyrAla	520
Db	1501	CGGATGACCTTCTGGCGGCTTCTGGCGGCTTCTGGCGGCTTCTGGCGGCTTCTGGCGG	1560

Qy 521 AlaAlaGlyGluLeuGluLysSerCysAspSerThrGlnAlaLeuValAlaPheSer 540
Db 1561 GCGCGGGAGCTGGAGAAGCTCTGTGATCTACCCAGGCGCTGTAGCTTCTCT 1620
Qy 541 AspValAspLeuLeuGluAlaSerGluAlaGlyArgProProGlyLeuGluThrTyr 560
Db 1621 GATGTGGATCTCAATCTGAAAGCTCTGAAAGCTGGGCGGCCCTTGGGCTGGAGACCTAT 1680
Qy 561 GlyPheProSerValThrLeuLeuSerCysGlnGlnProGlyAlaProArgLeuGluGly 580
Db 1681 GGCTTCCCTCAGTAGACCTCATCTCTGTGTCAGCAGCCAGGGGCCCCAGGCTGGAGGGC 1740
Qy 581 SerHisCysValGluProGluGlyAsnHisPheGlyAsnProGlnProSerMetAspGly 600
Db 1741 AGCCATTGTAGAGCCAGAGGGACCACTTTGGGAACCCCAACCTCTCATGTGATGA 1800
Qy 601 GluLeuLeuLeuArgAlaGluGlySerThrProAlaGlyGlyGlyLeuSerGlyGlyGly 620
Db 1801 GAACTGTCTGAGGCGAGAGGATCTACGCCAGCGGTGGAGGCTTGTACAGGGGTGGC 1860
Qy 621 GlyPheGlnProSerGlyLeuAlaPheAlaSerHisVal 633
Db 1861 GGCTTTAGCCCTCTGGCTTGGCCTTGTCTTCACACGTG 1899

RESULT 3

US-09-851-595-4
; Sequence 4, Application US/09851595
; Publication No. US20030166047A1
; GENERAL INFORMATION:
; APPLICANT: Gu, Wei
; TITLE OF INVENTION: NOVEL G-PROTEIN COUPLED RECEPTORS AND USES THEREFOR
; CURRENT APPLICATION NUMBER: US/09/851,595
; CURRENT FILING DATE: 2000-05-08
; PRIOR APPLICATION NUMBER: 09/556,588
; PRIOR FILING DATE: 2000-05-08
; PRIOR APPLICATION NUMBER: 60/132,896
; PRIOR FILING DATE: 1999-05-06
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: PatentIn ver. 2.0
; SEQ ID NO 4
; LENGTH: 2486
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (2)..(1900)
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (172)
; OTHER INFORMATION: n = any nucleotide
US-09-851-595-4

Alignment Scores:
Pred. No.: 0 Length: 2486
Score: 3305.00 Matches: 632
Percent Similarity: 99.84% Conservative: 0
Best Local Similarity: 99.84% Mismatches: 1
Query Match: 99.94% Indels: 0
DB: 10 Gaps: 0

US-10-664-667-5 (1-633) x US-09-851-595-4 (1-2486)

Qy 1 AsnThrThrHisTyrArgGluSerTrpTyrAlaCysArgTyrArgSerGlyIleProGly 20
Db 2 AATACGACTCACTATAGGAAAGCTGTGTCAGCTGCAGGTACCGGTCCGGAATTCGCGGG 61
Qy 21 SerThrHisAlaSerValGluArgSerGlnGlyLeuSerLeuProAlaHisProAlaSer 40
Db 62 TCGACCCAGCGGCTCGGTGGAGCGGAGGAGGCTGTAGCGCTGCGGCTCATCCAGCCTCT 121
Qy 41 LeuAlaAlaLeuAlaAlaSerAsnThrThrAlaSerGlyLysLeuGlu***AspThrPhe 60

Db 122 CTTGCTGCCCTAGCGGCTCCAAACACACCGCATCTGGGAAATTTGGAGCTNGACACCTTC 181
Qy 61 SerGlnLeuSerSerLeuGlnAlaLeuAspLeuSerTrpAlaIleAlaIleArgSerIleHis 80
Db 182 AGCCAGCTGAGCTCCCTGCAAGCCCTGGATCTTAGCTGGAAACGCATCCGGTCCATCCAC 241
Qy 81 ProGluAlaPheSerThrLeuHisSerLeuValLysLeuAspLeuThrAspAsnGlnLeu 100
Db 242 CTTGAGGCGCTTCTCCACCTTGCATCTCCCTGGTCAAGCTGGACCTGACAGACACACGCTG 301
Qy 101 ThrThrLeuProLeuAlaGlyLeuGlyLeuMetHisLeuLysLeuLysGlyAsnLeu 120
Db 302 ACCACACTGCCCTGGCTGGACTTGGGGCTTGTATGATCTTGAAGCTCAAAAGGAAACCTT 361
Qy 121 AlaLeuSerGlnAlaPheSerLysAspSerPheProLysLeuArgIleLeuGluValPro 140
Db 362 GCTCTCTCCAGGCGCTTCTCAAGGACAGTTTCCAAAACCTGAGGATCTCTGGAGGTGCT 421
Qy 141 TyrAlaTyrGlnCysCysProTyrGlyMetCysAlaSerPhePheLysAlaSerGlyGln 160
Db 422 TATGCTTACCAGTGTCTTCCCTATGGATGTGTGCCAGCTTCTCAAGGCTCTTGGGAG 481
Qy 161 TrpGluAlaGluAspLeuHisLeuAspAspGluSerSerLysArgProLeuGlyLeu 180
Db 482 TGGAGGCTGAAGACCTTCACTTGTATGATGAGGAGTCTTCAAAAAGGCCCTGGGCTC 541
Qy 181 LeuAlaArgGlnAlaGluAsnHisTyrAspGlnAspLeuAspGluLeuGlnLeuMet 200
Db 542 CTTGCCACAGACAGAGAACCACTATGACCAGGACCTGGATGAGCTCCAGCTGGAGATG 601
Qy 201 GluAspSerLysProHisProSerValGlnCysSerProThrProGlyProPheLysPro 220
Db 602 GAGGACTCAAAAGCCACACCCAGCTGTCCAGTGTAGCCCTACTCCAGGCCCTTCAAGCCC 661
Qy 221 CysGluTyrLeuPheGluSerTrpGlyIleArgLeuAlaValTrpAlaIleValLeu 240
Db 662 TGTGAGTACCTCTTTGAAAGCTGGGCGATCCGCTGGCCGTGGCGGCATCTGTTGCTC 721
Qy 241 SerValLeuCysAsnGlyLeuValLeuLeuThrValPheAlaGlyGlyProAlaProLeu 260
Db 722 TCCGTGCTCTGCAATGGAGCTGGTGTCTGCTGCTGCTGGCGGCTGCCCCCTG 781
Qy 261 ProProValLysPheValValGlyAlaIleAlaGlyAlaAsnThrLeuThrGlyIleSer 280
Db 782 CCCCCGTCAAGTTTGTGTAGGTGGATTGAGGCGCCACACCTTGTACTGGCATTTCC 841
Qy 281 CysGlyLeuLeuAlaSerValAspAlaLeuThrPheGlyGlnPheSerGluTyrGlyAla 300
Db 842 TGTGGCCTTCTAGCCTCAGTCGATGCCCTTGGTCACTTCTCTGAGTACGAGGCC 901
Qy 301 ArgTrpGluThrGlyLeuGlyCysArgAlaThrGlyPheLeuAlaValLeuGlySerGlu 320
Db 902 CGCTGGAGACGGGGCTAGGCTCCCGGCCACTTGGCTTCTTGGCAGTACTTGGGTCCGAG 961
Qy 321 AlaSerValLeuLeuLeuThrLeuAlaAlaValGlnCysSerValSerValSerCysVal 340
Db 962 GCATCGGTGCTGTCTCACTCTGGCGCGAGTGCAGTGCAGCGCTTCCGCTCTCTGTGTC 1021
Qy 341 ArgAlaTyrGlyLysSerProSerLeuGlySerValArgAlaGlyValLeuGlyCysLeu 360
Db 1022 CGGGCCTATGGGAAGTCCCCCTCCCTGGGCGAGGTTTCGAGCAGGGGTCTTAGGCTGCTG 1081
Qy 361 AlaLeuAlaGlyLeuAlaAlaLeuProLeuAlaSerValGlyGluTyrGlyAlaSer 380
Db 1082 GCATCGAGGGTGGCGCGCCACTGCCCTCGGCTCAGTGGGAGAAATACGGGGCTCC 1141
Qy 381 ProLeuCysLeuProTyrAlaProGluGlyGlnProAlaAlaLeuGlyPheThrVal 400
Db 1142 CCACTCTGCTGCGCTACGCCGCCACCTGAGGGTCAGCCAGCGCTTGGGCTTCCAGCTG 1201
Qy 401 AlaLeuValMetMetAsnSerPheCysPheLeuValValAlaGlyAlaTyrIleLysLeu 420
Db 1202 GCCCTGGTGTATGATGAACCTCTTCTGTTTCTGGTGGTGGCGGCTGCTACATCAAACTG 1261


```
Qy 421 TyrCysAspLeuProArgGlyAspPheGluAlaValTrpAspCysAlaMetValArgHis 440
Db 1262 TACTGTGACCTGCGCGGGGACATTTTCAGGCGCGTGTGGAGCTGGCCATGGTGAGGCAC 1321
Qy 441 VallalaTrpLeuIlePheAlaAspGlyLeuLeuTyrCysProValAlaPheLeuSerPhe 460
Db 1322 GTGGCCTGGCTCATCTTCGACAGCGGGCTCCTACTACTGTCCCGTGGCCCTTCTCAGCTTC 1381
Qy 461 AlaSerMetLeuGlyLeuPheProValTrpProGluAlaValLysSerValLeuLeuVal 480
Db 1382 GCCTTCATGCTGGGCTCTTCCCTGTACGCCCGGAGCGGCTCAAGTCGTCTGCTGGTG 1441
Qy 481 ValLeuProLeuProAlaCysLeuAsnProLeuLeuTyrLeuLeuPheAsnProHisPhe 500
Db 1442 GTGCTGCCCTGCTGCTGCTCAACCACTGCTGTACCTGCTCTTCAACCCCCACATTC 1501
Qy 501 ArgAspAspLeuArgArgLeuArgProArgAlaGlyAspSerGlyProLeuAlaTyrAla 520
Db 1502 CGGGATGACCTTCGGCGGCTTCGGCGCGCGCGAGGGGACTCAGGGCCCTAGCCTATGCT 1561
Qy 521 AlaAlaGlyGluLeuGluLysSerSerCysAspSerThrGlnAlaLeuValAlaPheSer 540
Db 1562 CGGCGGGGAGCTGGAGAGAGCTCCTGTGATCTACCCAGGCGCTGTAAGCTTCTCT 1621
Qy 541 AspValAspLeuIleLeuGluAlaSerGluAlaGlyArgProProGlyLeuGluThrTyr 560
Db 1622 GATGTGGATCTCATTTCTGAAGCTTCTGAAGCTGGCGGCGCCCTGCGGCTGGAGACCTAT 1681
Qy 561 GlyPheProSerValThrLeuIleSerCysGlnGlnProGlyAlaProArgLeuGluGly 580
Db 1682 GGCTTCCCTCAGTGACCTCATCTCTGTGTCAGCAGCCAGGGGCGCCCGAGGCTGGAGGC 1741
Qy 581 SerHisCysValGluProGluGlyAsnHisPheGlyAsnProGlnProSerMetAspGly 600
Db 1742 AGCCATTGTGTAGAGCCAGAGGGGACCCTTTGGGAAACCCCAACCCCTCCATGTATGGA 1801
Qy 601 GluLeuLeuLeuArgAlaGluGlySerThrProAlaGlyGlyGlyLeuSerGlyGlyGly 620
Db 1802 GAACCTGCTGTAGGCGCAGAGGGATCTACGCCAGCAGGTGGAGGCTTGTAGGGGGTGGC 1861
Qy 621 GlyPheGlnProSerGlyLeuAlaPheAlaSerHisVal 633
Db 1862 GGCTTTACGCCCTCTGGCTTGGCTTTGGCTTTTCACACGTG 1900

RESULT 4
US-10-664-667-4
; Sequence 4, Application US/10664667
; Publication No. US20040058377A1
; GENERAL INFORMATION:
; APPLICANT: Gu, Wei
; TITLE OF INVENTION: NOVEL G-PROTEIN COUPLED RECEPTORS AND USES THEREFOR
; FILE REFERENCE: MNI-080CP
; CURRENT APPLICATION NUMBER: US/10/664,667
; CURRENT FILING DATE: 2003-09-18
; PRIOR APPLICATION NUMBER: US/09/851,595
; PRIOR FILING DATE: 2000-05-08
; PRIOR APPLICATION NUMBER: 09/555,588
; PRIOR FILING DATE: 2000-05-08
; PRIOR APPLICATION NUMBER: 60/132,896
; PRIOR FILING DATE: 1999-05-06
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 4
; LENGTH: 2486
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (2)..(1900)
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (172)
```

OTHER INFORMATION: n = any nucleotide

US-10-664-667-4

```
Alignment Scores:
Pred. No.: 0 Length: 2486
Score: 3305.00 Matches: 632
Percent Similarity: 99.84% Conservative: 0
Best Local Similarity: 99.84% Mismatches: 1
Query Match: 99.94% Indels: 0
DB: 17 Gaps: 0
```

US-10-664-667-5 (1-633) x US-10-664-667-4 (1-2486)

```
Qy 1 AsnThrThrHisTyrArgGluSerTrpTyrAlaCysArgTyrArgSerGlyIleProGly 20
Db 2 AATAGCACTACTATAGGAAAGCTGGTAGCCCTGCAGGTACCGGTCCGGAATTCCTCCGGG 61
Qy 21 SerThrHisAlaSerValGluArgSerGlnGlyLeuSerLeuProAlaHisProAlaSer 40
Db 62 TCGACCCACGCGTCCGTGGAGCGGAGCCAGGGTCTGAGCCTGCCGGCTCATCCAGCCTCT 121
Qy 41 LeuAlaAlaLeuAlaAlaSerAsnThrThrAlaSerGlyLysLeuGlu**AspThrPhe 60
Db 122 CTTGCTGCCCTTAGCGGCTTCCAAACACACCACTCTGGGAAATTCGAGCTTGACACCTTC 181
Qy 61 SerGlnLeuSerSerLeuGlnAlaLeuAspLeuSerTrpAsnAlaIleArgSerIleHis 80
Db 182 AGCCAGCTGAGCTCCCTGCAAGCCCTGGATCTTAGCTGGAAACGCCATCCGCTCCATCCAC 241
Qy 81 ProGluAlaPheSerThrLeuHisSerLeuValLysLeuAspLeuThrAspAsnGlnLeu 100
Db 242 CCTGAGGCTTCTCCACCCCTGCACCTCCCTGCTGAAGCTGGAGCTGCACAGACAACAGCTG 301
Qy 101 ThrThrLeuProLeuAlaGlyLeuGlyLeuMetHisLeuLysLeuGlyAsnLeu 120
Db 302 ACCACACTGCCCTTGGCTGGACTTGGGGCTTGATGATCTGAAGCTCAAAGGGAACTT 361
Qy 121 AlaLeuSerGlnAlaPheSerLysAspSerPheProLysLeuArgIleLeuGluValPro 140
Db 362 GCTCTCTCCAGGCTTCTCCAGGACAGATTTTCCAAAACCTGAGGATCCTGGAGGTGCT 421
Qy 141 TyrAlaTyrGlnCysCysProTyrGlyMetCysAlaSerPhePheLeuAlaSerGlyGln 160
Db 422 TATGCTACCAAGTGTGTCCTATGGATGTGTGCCAGCTTCTTCAAGGCTCTGGGGAG 481
Qy 161 TrpGluAlaGluAspLeuHisLeuAspAspGluGluSerSerLysArgProLeuGlyLeu 180
Db 482 TGGGAGGCTGAAGACCTTCACTTGTATGATGAGGAGTCTTCAAAAAGGCCCTTGGGCTC 541
Qy 181 LeuAlaArgGlnAlaGluAsnHisTyrAspGlnAspLeuAspGluLeuGlnLeuGluMet 200
Db 542 CTTGCCAGACAAGCAGAGAACCACTATGACCAGGACCTGGATGAGCTCCAGCTGGAGATG 601
Qy 201 GluAspSerLysProHisProSerValGlnCysSerProThrProGlyProPheLysPro 220
Db 602 GAGGACTCAAAGCCACACCCAGTGTCCAGTGTAGCCCTACTCCAGGCGCCCTTCAAGGCC 661
Qy 221 CysGluTyrLeuPheGluSerTrpGlyLeuArgLeuAlaValTrpAlaIleValLeuLeu 240
Db 662 TGTGAGTACCTCTTTTGAAGAGCTGGGGCATCCGCTTGGCGGCTGAGGCTCGTGTGCTC 721
Qy 241 SerValLeuCysAsnGlyLeuValLeuLeuThrValPheAlaGlyGlyProAlaProLeu 260
Db 722 TCCGTGCTCTGCANTGGAGCTGGTGTCTGACCTGTTCCGTGGCGGCTGCCCTCCCTG 781
Qy 261 ProProValLysPheValValGlyAlaIleAlaGlyAlaAsnThrLeuThrGlyIleSer 280
Db 782 CCCCCTGCAAGTTTGTGGTAGGTGCGATTGCAGGCGCAACACACTTGACTGGCATTTCC 841
Qy 281 CysGlyLeuLeuAlaSerValAspAlaLeuThrPheGlyGlnPheSerGluTyrGlyAla 300
Db 842 TGTGGCTTCTAGCCTCAGTCGATGCCCTTGTGTCAGTTTGTGTCAGTGTCTCTGAGTAGGAGGCC 901
```

```
QY 301 ArgTrpGluThrGlyLeuGlyCysArgAlaThrGlyPheLeuAlaValLeuGlySerGlu 320
Db 902 CGCTGGAGACGGGCTAGGCTGCGCGCCACTGGCTTCTGGCAGTACTTGGGTGCGAG 961
QY 321 AlaSerValLeuLeuThrLeuAlaAlaValGlnCysSerValSerValSerCysVal 340
Db 962 GCATCGGTGCTGCTCACTTGGCCGACGTGCGAGTGCAGCGTCTCGCTCTCTGTGTC 1021
QY 341 ArgAlaTyrGlyLyseSerProSerLeuGlySerValArgAlaGlyValLeuGlyCysLeu 360
Db 1022 CGGGCTATGGAGTCCCTCCCTGGCGAGCGTTCAGCAGGGGTCTTAGGTGCGCTG 1081
QY 361 AlaLeuAlaGlyLeuAlaAlaLeuProLeuAlaSerValGlyGlyTyrGlyAlaSer 380
Db 1082 GCATGGCAGGGCTGGCGCGCACTGCGCCCTGGCTCAGTGGGAGATACGGGGCTTC 1141
QY 381 ProLeuCysLeuProTyrAlaProProGluGlyGlnProAlaAlaLeuGlyPheThrVal 400
Db 1142 CCATCTGCTGCTGCGCCACTGAGGCTCAGCCAGCAGCGCTGGGCTTCACCGTG 1201
QY 401 AlaLeuValMetMetAsnSerPheCysPheLeuValValAlaGlyAlaTyrIleValLeu 420
Db 1202 GCCCTGGTGGATGAGTCACTCTCTGTTCTGTCGTCGGCGGTGCTCATCAAACTG 1261
QY 421 TyrCysAspLeuProArgGlyAspPheGluAlaValTrpAspCysAlaMetValArgHis 440
Db 1262 TACTGTGACCTGCGCGGGCGACTTTGAGGCGGTGTGGACTCGGCCATGTTGAGGCAC 1321
QY 441 VallalaTrpLeuIlePheAlaAspGlyLeuLeuTyrCysProValAlaPheLeuSerPhe 460
Db 1322 GTGGCTGGCTCATCTTCGCAGAGGGCTCTCTACTGTCCGTCGGCTTCTCAGCTTC 1381
QY 461 AlaSerMetLeuGlyLeuPheProValThrProGluAlaValLyseSerValLeuVal 480
Db 1382 GCCTCCATGCTGGGCTCTTCCCTGTACGCGCCGAGGCGCGCAAGTCTGTCTCTGCTGTG 1441
QY 481 ValLeuProLeuProAlaCysLeuAsnProLeuLeuTyrLeuLeuPheAsnProHisPhe 500
Db 1442 GTGCTGCCCTGCTGCTGCTCAACCATGCTGTGTACTGCTTTCACCCCACTTC 1501
QY 501 ArgAspAspLeuArgArgLeuArgProArgAlaGlyAspSerGlyProLeuAlaTyrAla 520
Db 1502 CGGGATGACCTTCGGCGGCTTCGGCGCGCGCGCAGGAGACTCAGGGCGCTAGCCTATGCT 1561
QY 521 AlaAlaGlyGluLeuGlyLyseSerCysAspSerThrGlnAlaLeuValAlaPheSer 540
Db 1562 CGCGCGGGAGCTGGAGAGAGCTCTGTGATTTCTACCCAGGCGCTGGTAGCCTTCTCT 1621
QY 541 AspValAspLeuIleLeuGluAlaSerGluAlaGlyArgProProGlyLeuGluThrTyr 560
Db 1622 GATGTGGATCTCATCTTGGAGCTTCTGAAGCTGGGCGGCGCCCTGGGCTGGAGCCTAT 1681
QY 561 GlyPheProSerValThrLeuIleSerCysGlnProGlyAlaProArgLeuGluGly 580
Db 1682 GGCTTCCCTCAGTGACCTCATCTCTGTCTCAGCAGCAGGGGCGCCAGGCTGGAGGGC 1741
QY 581 SerHisCysValGluProGluGlyAsnHisPheGlyAsnProGlnProSerMetAspGly 600
Db 1742 AGCCATTTGTGTAGAGCCAGAGGGGAACCACTTTGGGAACCCCAACCTCCATGGATGGA 1801
QY 601 GluLeuLeuArgAlaGluGlySerThrProAlaGlyGlyGlyLeuSerGlyGlyGly 620
Db 1802 GAATCTGTGTGAGGGCAGAGGGATCTTAGCCAGCAGGTGGAGGCTTGTAGGGGGTGGC 1861
QY 621 GlyPheGlnProSerGlyLeuAlaPheAlaSerHisVal 633
Db 1862 GGCTTTGAGCCCTCTGGCTGGCCCTTGTGCTTACACGTG 1900
```

RESULT 5

US-09-851-595-9

; Sequence 9, Application US/09851595

; Publication No. US20030166047A1

; GENERAL INFORMATION:

```
; APPLICANT: Gu, Wei
; TITLE OF INVENTION: NOVEL G-PROTEIN COUPLED RECEPTORS AND USES THEREFOR
; FILE REFERENCE: MNI-080CP
; CURRENT APPLICATION NUMBER: US/09/851,595
; CURRENT FILING DATE: 2000-05-08
; PRIOR APPLICATION NUMBER: 09/556,588
; PRIOR FILING DATE: 2000-05-08
; PRIOR APPLICATION NUMBER: 60/132,896
; PRIOR FILING DATE: 1999-05-06
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 9
; LENGTH: 2208
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(2208)
US-09-851-595-9
Alignment Scores:
Pred. No.: 3,08e-308 Length: 2208
Score: 3045.50 Matches: 602
Percent Similarity: 92.61% Conservative: 12
Best Local Similarity: 90.80% Mismatches: 19
Query Match: 92.09% Indels: 32
DB: 10 Gaps: 4
```

US-10-664-667-5 (1-613) x US-09-851-595-9 (1-2208)

```
QY 1 AsnThrThrHisTyr-----CysArgTyrArg 15
Db 224 AACTCCACACACTATCTCTGAATGGTCCATGCAGATCCAGAGTTTCCAGATCTCAAG 283
QY 6 ---ArgGluSerTrpTyrAla-----CysArgTyrArg 15
Db 284 GCACCACACGCTGGAGATCTTGACCTCGACCGCGCAGCATCCGGCTGTCCCATCGG 343
QY 16 -----SerGlyIleProGlySer-ThrHisAlaSerValGluArgSerGln---G1 31
Db 344 GGATGTGCCAACAGCTGCCCAGGCTCCGAGTCTCTGGAACTGTCTCACAATCAAATTGAGG 403
QY 31 YLeuSerLeuProAlaHisProAlaSerLeuAlaLeuAlaAlaSerAsnThrThrAl 51
Db 404 AGTGTGCCACCTTGACAGGTGTCAGNAATTTGAGGAGAA--TCGGCTTCCAAACACACCGC 462
QY 51 aSerGlyLyLeuGlu***AspThrPheSerGlnLeuSerSerLeuGlnAlaLeuAspLe 71
Db 463 ATCTGGGAAATTGGAGCT--GACACCTTCAGCCAGCTGAGCTCCCTGCAAGCCCTGGATCT 521
QY 71 uSerTrpAsnAlaIleArgSerIleHisProGluAlaPheSerThrLeuHisSerLeuVa 91
Db 522 TAGCTGGAGCCATCCGGTCCATCCACCTTGAGGCTTCTCCACCCCTGCACCTCCCTGGT 581
QY 91 lLysLeuAspLeuThrAspAsnGlnLeuThrThrLeuProLeuAlaGlyLeuGlyGly 111
Db 582 CAAGCTGGAGCTGACAGAACACAGCTGACCACTGCCCTCCCTGGCTGGAGCTTGGGGCTT 641
QY 111 uMetHisLeuLyLeuLyGlyAsnLeuAlaLeuSerGlnAlaPheSerLyAspSerPh 131
Db 642 GATGCATCTGAAGCTCAAAGGGAACCTTGTCTCTCCAGGCTTCTTCCAGGACAGTCTT 701
QY 131 eProLyLeuArgIleLeuGluValProTyrAlaTyrGlnCysCy8ProTyrGlyMetCy 151
Db 702 CCCAAACCTGAGGATCTCGAGGTGCTTATGCTTACCACTGCTGCTCCCTATGGGATGTG 761
QY 151 sAlaSerPhePheLyAlaSerGlyGlnTrpGluAlaGluAspLeuHisLeuAspG1 171
Db 762 TGCAGCTTCTTCAAGGCTCTCGGAGGTGGAGGCTGAAGACCTTCACTTGAATGATGA 821
QY 171 uGluSerSerLyArgProLeuGlyLeuAlaArgGlnAlaGluAsnHisTyrAspG1 191
Db 822 GGAGTCTTCAAAAAGGCCCTTGGGCTTCTTGGCAGACAGAGAGAACCACTATGACCA 881
```

Qy	191	nAspLeuApGluLeuGlnLeuGluMetGluAspSerIysProHisProSerValGlnCy	211
Db	882	GGACCTGGATGAGCTCCAGCTGGAGATGGAGGACTCAAAGCCACACCCCACTGTCCAGTG	941
Qy	211	sSerProThrProGlyProPheIysProCysGluTyrlleuPheGluSerTrpGlyIleAr	231
Db	942	TAGCCCTACTCCAGGCCCTTCAAGCCCTGTGAGTACTCTTTTGAAGCTGGGGCATCCG	1001
Qy	231	gLeuAlaValTrpAlaIleValLeuLeuSerValLeuCysAsnGlyLeuValLeuLeuTh	251
Db	1002	CCTGGCCGTGTGGGCCATCGTGTGTCTCTCGTGCTCTGCAATGGACTGGTGTCTGTGAC	1061
Qy	251	rValPheAlaGlyClyProAlaProLeuProLeuProValIysPheValValGlyAlaIleAl	271
Db	1062	CGTGTTCGTGGCGGGCTGCCCCCTCGCCCGGTCAAGTTTGTGTGGTGGGATGCG	1121
Qy	271	aGlyAlaAsnThrLeuThrGlyIleSerCysGlyLeuLeuAlaSerValAspAlaLeuTh	291
Db	1122	AGGCGCCACACCTTGACTGGCATTTCTGTGGCTTCTAGCCTCAGTCGATGCCCTGAC	1181
Qy	291	rPheGlyGlnPheSerGluTyrlGlyAlaArgTrpGluThrGlyLeuGlyCysArgAlaTh	311
Db	1182	CTTTGGTCAGTTCTCTGAGTACGGAGCCCGCTGGGAGACGGGCTAGGCTGCCGGCCAC	1241
Qy	311	rGlyPheLeuAlaValLeuGlySerGluAlaSerValLeuLeuLeuThrLeuAlaIaVa	331
Db	1242	TGGCTTCTTGGCAGTACTTGGGTCCGAGGCCATCGGTGCTGCTGCTACTCTGGCCGAGT	1301
Qy	331	lGlnCysSerValSerValSerCysValArgAlaTyrlGlyIysSerProSerLeuGlyISe	351
Db	1302	GCAGTGACGGCTCTCCGTCTCTGTGTCCGGCCCTATGGGAAGTCCCCCTCCCTGGGAG	1361
Qy	351	rValArgAlaGlyValLeuGlyCysLeuAlaLeuAlaGlyLeuAlaAlaLeuProLe	371
Db	1362	CGTTCGACGAGGGTCTTAGGCTGCCCTGGCACTGGCAGGGGTGGCCCGCAGCTGCCCT	1421
Qy	371	uAlaSerValGlyGluTyrlGlyAlaSerProLeuCysLeuProTyrlAlaProProGluGl	391
Db	1422	GGCCTCAGTGGGAGAAATACGGGGCTCCCCACTCTGGCTGGCCCTACCGCCACCTGAGG	1481
Qy	391	YGlnProAlaAlaLeuGlyPheThrValAlaLeuValMetMetAsnSerPheCysPheLe	411
Db	1482	TCAGCCAGCAGCCCTGGGCTTCACCGTGGCCCTGGTGATGATGAACCTCTCTCTTTCT	1541
Qy	411	uValValaGlyAlaTyrlleIysLeuTyrlCysAspLeuProArgGlyAspPheGluAl	431
Db	1542	GGTCGTGGCCGTGCCATACATCAAACTGTACTGTGACTCTGGCGGGGGCGACTTTGAGGC	1601
Qy	431	aValTrpAspCysAlaMetValArgHisValAlaTrpIlePheAlaAspGlyIleuLe	451
Db	1602	CGTGTGGACTTGGCCCATGTGTGAGGACGTGGCTTGGCTCATCTTCGACAGCGGGCTCT	1661
Qy	451	uTyrlCysProValAlaPheLeuSerPheAlaSerMetLeuGlyLeuPheProValThrPr	471
Db	1662	CTACTGTCCGTGGCTTCTCAGTTCGCCTCCATGCTGGCCCTCTTCCCTGTTCAGCC	1721
Qy	471	oGluAlaValIysSerValLeuLeuValValLeuProLeuProAlaCysLeuAsnProLe	491
Db	1722	CGAGGCCGTCAAGTCTGTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT	1781
Qy	491	uLeuTyrlleuLeuPheAsnProHisPheArgAspLeuArgGluArgProArgAl	511
Db	1782	GCTGTACTGTCTTTAAACCCCCACTTCCGGGATGACCTTCGGCGGGCTTCGGCCCGGC	1841
Qy	511	aGlyAspSerGlyProLeuAlaTyrlAlaAlaGlyIleGluLeuGlyIysSerSerCysAs	531
Db	1842	AGGGGACTCAGGGCCCCCTAGCCTATGCTGTGGCCCGGGAGCTGGAGAAGAGCTCCTGTGA	1901
Qy	531	pSerThrGlnAlaLeuValAlaPheSerAspValAspIleIleLeuGluAlaSerGluAl	551
Db	1902	TTCTACCCAGCCCTGTGTAGCTTCTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT	1961

[illegible]

```
Qy 51 aserGlyLysLeuGlu***AspThrPheSerGlnLeuSerSerLeuGlnAlaLeuAspLe 71
Db 463 ATCTGGGAAATTGGAGCT-GACACCTTCAGCCAGCTGAGCTCCCTGCGAGCCCTGGATCT 521
Qy 71 uSerTrpAsnAlaIleArgSerIleHisProGluAlaPheSerThrLeuHisSerLeuVa 91
Db 522 TAGCTGGAAACCCATCCGGTCCATCCAGCCCTGAGGCCCTTCTCCACCCCTGCACCTCCCTGGT 581
Qy 91 lLysLeuAspLeuThrAspAsnGlnLeuThrThrLeuProLeuAlaGlyLeuGlyLe 111
Db 582 CAACTGGACCTGACAGACAAACAGCTGACACACTGCCCCCTGGCTGGACTTGGGGGCTT 641
Qy 111 uMetHisLeuLysLeuLysGlyAsnLeuAlaLeuSerGlnAlaPheSerLysAspSerPh 131
Db 642 GATGCATCTGAAGCTCAAGAGGAACCTTGCTCTCTCCAGGCCCTTCTCCAGGACAGTTT 701
Qy 131 eProLysLeuArgIleLeuGluValProTyrAlaTyrGlnCysCysProTyrGlyMetCy 151
Db 702 CCCAAAACCTGAGGATCCTGGAGGTGCCCTTATGCCCTACCAGTGCTGCTCCCTATGGATGTG 761
Qy 151 sAlaSerPhePheLysAlaSerGlyGlnTrpGluAlaGluAspLeuHisLeuAspAspG 171
Db 762 TGCCAGCTTCTTCAAGGCCCTCTGGGCAGGTGGAGGCTGAAGACCTTCACCTTGATGATGA 821
Qy 171 uGluSerSerLysArgProLeuGlyLeuLeuAlaArgGlnAlaGluAsnHisTyrAspG 191
Db 822 GGAGTCTTTCAANAAGGCCCTGGGCCCTCTTGGCCAGACACAGACAGAAACCACTATGACCA 881
Qy 191 nAspLeuAspGluLeuGlnLeuGluMetGluAspSerLysProHisProSerValGlnCy 211
Db 882 GGACCTGATGAGCTCCAGCTGGAGATGGAGGACTCAAAAGCCACACCCAGTGCAGTG 941
Qy 211 sSerProThrProGlyProPheLysProCysGluTyrIleuPheGluSerTrpGlyIleAr 231
Db 942 TAGCCCTACTCCAGGCCCTTCAAGGCCCTGTGAGTACCTCTTTTGAAGCTGGGGCATCCG 1001
Qy 231 gLeuAlaValTrpAlaIleValLeuLeuSerValLeuCysAsnGlyLeuValLeuLeuTh 251
Db 1002 CCTGGCCGTGTGGGCCATCGHTTGTCTCCGTGCTCTGCANATGACTGTGTGCTGTGAC 1061
Qy 251 rValPheAlaGlyGlyProAlaProLeuProProValLysPheValValGlyAlaIleAl 271
Db 1062 CGTGTCTGCTGGCGGCCCTGCCCCCTGCCCCCGTCAAGTTTGTGTAGTGCGATTGC 1121
Qy 271 aGlyAlaAsnThrLeuThrGlyIleSerCysGlyLeuLeuAlaSerValAspAlaLeuTh 291
Db 1122 AGGCGCCAAACACCTTGACTGGCAITTCCTGTGGCTCTTAGCCCTCAGTCGATGCCCTGAC 1181
Qy 291 rPheGlyGlnPheSerGluTyrGlyAlaAtgTrpGluThrGlyLeuGlyCysArgAlaTh 311
Db 1182 CTTTGGTCAGTTCTCTGAGTACGAGCCCGCTGGAGACGGGGCTAGGCTGCCCGGCCAC 1241
Qy 311 rGlyPheLeuAlaValLeuGlySerGluAlaSerValLeuLeuLeuThrLeuAlaIleAlaVa 331
Db 1242 TGGCTTCTTGGCAGTACTTGGGTGGAGGATCGTGTCTGTCTCCTCTGCGCGCAGT 1301
Qy 331 lGlnCysSerValSerValSerCysValArgAlaTyrGlyLysSerProSerLeuGlySe 351
Db 1302 GCAGTGCAGCGCTCTCCGTCTCTGTGTCCGGGCCCTATGGGAAGTCCGCCCTCCCTGGGCAG 1361
Qy 351 rValArgAlaGlyValLeuGlyCysLeuAlaLeuAlaGlyLeuAlaIleAlaLeuProLe 371
Db 1362 CGTTTCGAGCAGGGGTCTTAGGCTCCCTGGCACTGGCAGGGCTGGCCCGCCAGCTGCCCT 1421
Qy 371 uAlaSerValGlyGluTyrGlyAlaSerProLeuCysLeuProTyrAlaProProGluGl 391
Db 1422 GGCCTCAGTGGGAATACGGGGCTTCCCACTCTGCTGCCCTACGCGCCACCTGAGGG 1481
Qy 391 yGlnProAlaAlaLeuGlyPheThrValAlaLeuValMetMetAsnSerPheCysPheLe 411
Db 1482 TCAGCCAGCAGCCCTGGGCTTCAACCGTGGCCCTGTGATGATGAACCTCTCTGTTTCT 1541
Qy 411 uValValAlaGlyAlaTyrIleLysLeuTyrCysAspLeuProArgGlyAspPheGluAl 431
```

```
Db 1542 GGTCTGTCGGCGGTTCCTACATCAAACTGTACTGTGACCTGCGCGGGCGAGCTTTGAGGC 1601
Qy 431 aValTrpAspCysAlaMetValArgHisValAlaTrpLeuIlePheAlaAspGlyLeuLe 451
Db 1602 CGTGTGGGACTTGGCCATGGTGGAGCACGTGGCTGGCTCATCTTCGACAGCGGGCTCCT 1661
Qy 451 uTyrCysProValAlaPheLeuSerPheAlaSerMetLeuGlyLeuPheProValThrPr 471
Db 1662 CTACTGTCCCGTGGCTTCTCTCAGCTTCGCTTCATGCTGGGCTCTTCCCTGTACGCC 1721
Qy 471 oGluAlaValLysSerValLeuLeuValValLeuProLeuProAlaCysLeuAsnProLe 491
Db 1722 CGAGGCCGTCAAGTCTGTCTGTGTGGTGTGCTGCCCTGCCTGCCTCAACCCACT 1781
Qy 491 uLeuTyrLeuLeuPheAsnProHisPheArgAspAspLeuArgArgLeuArgProArgAl 511
Db 1782 GCTGTACTCTCTTCAACCCCACTTCCGGGATGACCTTCGGCGGCTTCGGCCCGCCGC 1841
Qy 511 aGlyAspSerGlyProLeuAlaTyrAlaAlaGlyGluLeuGluLysSerSerCysAs 531
Db 1842 AGGGGACTCAGGGGCCCTAGCCTATGCTGCGCCGGGAGCTGGAGAGAGCTCCTGTGA 1901
Qy 531 pSerThrGlnAlaLeuValAlaPheSerAspValAspLeuIleLeuGluAlaSerGluAl 551
Db 1902 TTCTACCCAGGCCCTGGTAGCCTTCTGTGTGGATCTCATCTTGGAGCTTCTGAAGC 1961
Qy 551 aGlyArgProProGlyLeuGluThrTyrGlyPheProSerValThrLeuLysSerCysGl 571
Db 1962 TGGCGCGCCCTGGGCTGGAGACCTATGGCTTCCCTCAGTGACCCCTCATCTCTGTCA 2021
Qy 571 nGlnProGlyAlaProArgLeuGlySerHisCysValGluProGluGlyAsnHisPhe 591
Db 2022 GCAGCCAGGGGCCCCAGGCTGGAGGACGCACTGTGTAGAGCCAGAGGGAACCACTT 2081
Qy 591 eGlyAsnProGlnProSerMetAspGlyGlyLeuLeuLeuLeuArgAlaGluGlySerThrPr 611
Db 2082 TGGGAACCCCAACCCCTCCATGGATGGAGAACTGCTGCTGAGGGCAGAGGATCTAGCC 2141
Qy 611 oAlaGlyGlyGlyLeuSerGlyGlyGlyPheGlnProSerGlyLeuAlaPheAlaSe 631
Db 2142 AGCAGGTGGAGGCTTGTTCAGGGGTGGCGGCTTTCAGCCCTCTGGCTTGGCTTGTCTT 2201
Qy 631 rHisVal 633
Db 2202 ACACGTG 2208

RESULT 7
US-09-851-595-7
; Sequence 7, Application US/09851595
; Publication No. US20030166047A1
; GENERAL INFORMATION:
; APPLICANT: Gu, Wei
; TITLE OF INVENTION: NOVEL G-PROTEIN COUPLED RECEPTORS AND USES THEREFOR
; FILE REFERENCES: MNI-080CP
; CURRENT APPLICATION NUMBER: US/09/851,595
; PRIOR FILING DATE: 2000-05-08
; PRIOR APPLICATION NUMBER: 09/556,588
; PRIOR FILING DATE: 2000-05-08
; PRIOR APPLICATION NUMBER: 60/132,896
; PRIOR FILING DATE: 1999-05-06
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 7
; LENGTH: 2711
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(2208)
US-09-851-595-7

Alignment Scores:
```

Pred. No.:	4,126-308	Length:	2711
Score:	3045.50	Matches:	602
Percent Similarity:	92.61%	Conservative:	12
Best Local Similarity:	90.80%	Mismatches:	19
Query Match:	92.03%	Indels:	32
DB:	10	Gaps:	4

US-10-664-667-5 (1-633) x US-09-851-595-7 (1-2711)

QY	1	AenThrThrHisTyr-----	5
DB	224	AACTCCACACATATCTCTGAATGTGGCCATGGACATCCAGAGATTTCAGATCTCAAG	283
QY	6	---ArgGluSerTyrTyrAla-----	15
DB	284	GCACACACAGCCTGGAGATCTCCACCTGCACCCCGCAGGCATCCGGCTGCCATCGG	343
QY	16	-----SerGlyLeuProGlySer-ThrHisAlaSerValGluArgSerGln---	31
DB	344	GGATGTGCCAACAGCTGCCAGGCTCCGAGTCTCCGAACTCTCTCACAATCAAAATTGAGG	403
QY	31	YLeuSerLeuProAlaHisProAlaSerLeuAlaAlaLeuAlaAlaSerAsnThrAl	51
DB	404	AGCTGCCAGCCTGCACAGGTGCAGAAATTGGAGAAA-TCGGCCTCCACACACACCGC	462
QY	51	AserGlyLysLeuGlu***AspThrPheSerGlnLeuSerSerLeuGlnAlaLeuAspLe	71
DB	463	ATCTGGAAATTTGGAGCT-GACACCTTCAGCCAGCTGAGCTCCCTGCAAGCCCTGATCT	521
QY	71	UserTrpAsnAlaIleArgSerIleHisProGluAlaPheSerThrLeuHisSerLeuVa	91
DB	522	TAGCTGGAAACGCATCCGCTCCATCCACCTGAGGCTTCTCCACCTGCACCTCCCTGGT	581
QY	91	LysLeuAspLeuThrAspAsnGlnLeuThrThrLeuProLeuAlaGlyLeuGlyLe	111
DB	582	CAAGCTGACCTGCACAGAACACAGCTGACCACTGCCCTTGGCTGGCTGGGGGCTT	641
QY	111	uMethHisLeuLysLeuLysGlyAsnLeuAlaLeuSerGlnAlaPheSerLysAspSerPh	131
DB	642	GATGCATCTGAAGCTCAAGGGAACCTTCTCTCTCCAGGCTTCTCCAGGACAGTTT	701
QY	131	eProLysLeuArgIleLeuGluValProTyrAlaTyrGlnCysCysProTyrGlyMetCy	151
DB	702	CCCAAACTCAGGATCTCGAGGTGCCTTATGCCTACCACTGCTGCTCCCTATGGGATGTG	761
QY	151	eAlaSerPhePheLysAlaSerGlyGlnTrpGluAlaGluAspLeuHisLeuAspG1	171
DB	762	TGCCAGCTTCTTCAAGGCTCTGGGAGGTGGAGGCTGAAGACCTTCACCTTTGATGATGA	821
QY	171	uGluSerSerLysArgProLeuGlyLeuLeuAlaArgGlnAlaGluAsnHisTyrAspG1	191
DB	822	GGAGTCTTCAAAAGGCCCTCGGGCTCTTGGCAGACAGAGAGAACCACTATGACCA	881
QY	191	nAspLeuAspGluLeuGlnLeuGluMetGluAspSerLysProHisProSerValGlnCy	211
DB	882	GGACCTGGATGAGCTCCAGCTGGAGATGGAGACTCAAAAGCCACACCCAGTGTCCAGTG	941
QY	211	sSerProThrProGlyProPheLysProCysGluTyrLeuPheGluSerTrpGlyLeaR	231
DB	942	TAGCCCTACTCCAGGCCCTTCAAGCCCTGTGAGTACCTCTTTGAAAGCTGGGGGATCCG	1001
QY	231	gLeuAlaValTrpAlaIleValLeuSerValLeuCysAsnGlyLeuValLeuLeuTh	251
DB	1002	CCTGGCCGTGGGGCCATCGTGTCTCTCCGTGCTCTGCAATGACATGGTGTGCTGCTGAC	1061
QY	251	rValPheAlaGlyGlyProAlaProLeuProProValLysPheValValGlyAlaIleAl	271
DB	1062	CGTGTCTGGCGGCTGCCCCCTGCCCCCGGTCAAGTTTGTGGTAGGTGCGATTGC	1121
QY	271	aglyAlaAsnThrLeuThrGlyLysSerCysGlyLeuLeuAlaSerValAspAlaLeuTh	291
DB	1122	AGGCCCAACACCTTTGACTGGCATTTCTGTGGCCCTTCTAGCCTCAGTCGATGCCCTGAC	1181

QY	291	rPheGlyGlnPheSerGluTyrGlyAlaArgTrpGluThrGlyLeuGlyCysArgAlaTh	311
DB	1182	CTTTGGTCAAGTCTCTGAGTACGGAGCCCGCTGGAGACGGGCTAGGCTGCGGGCCAC	1241
QY	311	rGlyPheLeuAlaValLeuGlySerGluAlaSerValLeuLeuLeuThrLeuAlaAlaVa	331
DB	1242	TGGCTTCTTGGCAGTACTTGGGTGGAGGCATCGGTGCTGCTGCTCACTCTGGCGCGAGT	1301
QY	331	lGlnCysSerValSerValSerCysValArgAlaTyrGlyLysSerProSerLeuGlySe	351
DB	1302	GCAGTGCAGCTCTCCGTCTCTGTGTCGGGCTATGGGAAGTCCCTCTCCCTGGGAG	1361
QY	351	rValArgAlaGlyValLeuGlyCysLeuAlaLeuAlaGlyLeuAlaAlaLeuProLe	371
DB	1362	CGTTCCAGCAGGGGTCTAGGCTGCTGGCACTGGCAGGGGTGGCGCGCAGCTCCCT	1421
QY	371	uAlaSerValGlyGluTyrGlyAlaSerProLeuCysLeuProTyrAlaProProGluG1	391
DB	1422	GGCTTCAGTGGAGAAATACGGGGCTCCCTCTGCTGCTGCTACCGCGCACCTGAGGG	1481
QY	391	yGlnProAlaAlaLeuGlyPheThrValAlaLeuValMetMetAsnSerPheCysPheLe	411
DB	1482	TCAGCCAGCAGCCCTGGGCTTCCCGTGGCTTGGTGTATGATGAATCTCTCTGTCTCT	1541
QY	411	uValValAlaGlyAlaTyrIleLysLeuTyrCysAspLeuProArgGlyAspPheGluAl	431
DB	1542	GGTGGTGGCTGCTTACATCAAACTGTACTGTGACCTGGCGGGGCGAGCTTTGAGGC	1601
QY	431	avalTrpAspCysAlaMetValArgHisValAlaTrpLeuIlePheAlaAspGlyLeuLe	451
DB	1602	CGTGGGACTTGCCTTGGTGGAGCACGTGGCTTGGCTCATCTTCGCGAGACGGGCTCT	1661
QY	451	uTyrCysProValAlaPheLeuSerPheAlaSerMetLeuGlyLeuPheProValThrPr	471
DB	1662	CTACTGCTCCGTGGCTTCTCAGCTTCCCTCCATGCTGGGCTCTTCCCTGTCCAGGC	1721
QY	471	oGluAlaValLysSerValLeuLeuValValLeuProLeuProAlaCysLeuAsnProLe	491
DB	1722	CGAGGCCGTCAAGTCTGCTGCTGGTGGTGGCTGGCTGGCTGGCTGGCTGGCTGGCT	1781
QY	491	uLeuTyrLeuLeuPheAsnProHisPheArgAspAspLeuArgLeuArgProArgAl	511
DB	1782	GCTGTACTGCTTCAACCCCTTCCGGGATGACCTTCGGGGCTTCGGCCCCCGGC	1841
QY	511	aglyAspSerGlyProLeuAlaTyrAlaAlaGlyGluLeuGluLysSerSerCysAs	531
DB	1842	AGGGAGCTCAGGGCCCTAGGCTATGCTGGCGCGGGAGCTGGAGAGAGAGCTCTCTGA	1901
QY	531	pSerThrGlnAlaLeuValAlaPheSerAspValAspLeuIleLeuGluAlaSerGluAl	551
DB	1902	TTCTACCCAGGCCCTGGTGGTCTCTCTGATGTGATCTCATTTCTGGAAGCTTTCTGAAGC	1961
QY	551	aglyArgProProGlyLeuGluThrTyrGlyPheProSerValThrLeuIleSerCysG1	571
DB	1962	TGGCGGGCCCTGGGCTGGAGACTATGGCTTCCCTCAGTGACCTCATCTCTCTGCA	2021
QY	571	nGlnProGlyAlaProArgLeuGluLysSerHisCysValGluProGluGlyAsnHisPh	591
DB	2022	GCAGCGGGGCCCTCCAGGCTGGAGGCGAGCATTTGTAGAGCCAGAGGGGAACACTT	2081
QY	591	eGlyAsnProGlnProSerMetAspGlyGluLeuLeuLeuArgAlaGluGlySerThrPr	611
DB	2082	TGGGAACCCCAACCTCCATGGATGGAGAACTGCTGCTGAGGGCAGAGGGATCTACGCC	2141
QY	611	oAlaGlyGlyGlyLeuSerGlyGlyGlyPheGlnProSerGlyLeuAlaPheAlaSe	631
DB	2142	AGCAGTGGAGGCTTGTGAGGGGTGGCGGCTTTGAGCCCTTGGCTTGGCTTGGCTTTC	2201
QY	631	rHisVal 633	
DB	2202	ACAGTG 2208	

RESULT 8

```
US-10-664-667-7
; Sequence 7, Application US/10664667
; Publication No. US20040058377A1
; GENERAL INFORMATION:
; APPLICANT: Gu, Wei
; TITLE OF INVENTION: NOVEL G-PROTEIN COUPLED RECEPTORS AND USES THEREFOR
; FILE REFERENCE: MNI-080CP
; CURRENT APPLICATION NUMBER: US/10/664,667
; CURRENT FILING DATE: 2003-09-18
; PRIOR FILING DATE: 2000-05-08
; PRIOR APPLICATION NUMBER: US/09/851,595
; PRIOR FILING DATE: 2000-05-08
; PRIOR APPLICATION NUMBER: 09/556,588
; PRIOR FILING DATE: 2000-05-08
; PRIOR APPLICATION NUMBER: 60/132,896
; PRIOR FILING DATE: 1999-05-06
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 7
; LENGTH: 2711
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(2208)
US-10-664-667-7
Alignment Scores:
Pred. No.: 4,12e-308 Length: 2711
Score: 3045.50 Matches: 602
Percent Similarity: 92.61% Conservative: 12
Best Local Similarity: 90.80% Mismatches: 19
Query Match: 92.09% Indels: 32
DB: 17 Gaps: 4

US-10-664-667-5 (1-633) x US-10-664-667-7 (1-2711)
QY 1 AsnThrThrHisTyr----- 5
DB 224 AACTCCACACACTATCTCTGAATGGTGCATGGACATCCAGGAGTTTCCAGATCTCAAG 283
QY 6 ---ArgGluSerTrpTyrAla-----CysArgTyrArg 15
DB 284 GCACACCAGCCTGGAGATCTGACCTCCACCCGCGCAGGCATCCGGTGTCTCCCATCGG 343
QY 16 -----SerGlyIleProGlySer-ThrHisAlaSerValGluArgSerGln-----G1 31
DB 344 GGATGTGCCAACAGCTGCCCGAGGCTCCGAGTCTCTGGAACTGTCTCACAATCAAAATTGAGG 403
QY 31 YLeuSerLeuProAlaHisProAlaSerLeuAlaAlaLeuAlaAlaSerAsnThrThrAl 51
DB 404 AGCTGCCAGCCTGCACAGGTGTGAGAAATTGGAGGAAA-TCGGCCTCCAAACACCCGC 462
QY 51 aSerGlyLysLeuGlu***AspThrPheSerGlnLeuSerSerLeuGlnAlaLeuAspLe 71
DB 463 ATCTGGAAATTGGAGCT-GACACCTTCAGCCAGCTGAGCTCCCTGCAAGCCCTGGATCT 521
QY 71 uSerTrpAsnAlaIleArgSerIleHisProGluAlaPheSerThrLeuHisSerLeuVa 91
DB 522 TAGCTGGAACCCCATCCGGTCCATCCACCTCGAGGCTTCTCCACCTGCACTCCCTCGT 581
QY 91 lLysLeuAspLeuThrAspAsnGlnLeuThrLeuProLeuAlaGlyLeuGlyLe 111
DB 582 CAAGCTGGACCTGACAGAACACAGCTGACCACTGCCCTTGGCTGGACTTGGGGGCTT 641
QY 111 uMetHisLeuLysLeuLysGlyAsnLeuAlaLeuSerGlnAlaPheSerLysAspSerPh 131
DB 642 GATGCATCTGAAGCTCAAAAGGGAACCTTGTCTCTCTCCAGGCCCTTCTCCAAGGACAGTTT 701
QY 131 eProLysLeuArgIleLeuGluValProTyrAlaTyrGlnCysCysProTyrGlyMetCy 151
DB 702 CCAAAAACCTGAGGATCTGGAGGTGCCTTATGCTTACCACTGCTGCTGCTATGGGGATGTG 761
QY 151 sAlaSerPhePheLysAlaSerGlyGlnTrpGluAlaGluAspLeuHisLeuAspAspG1 171
DB 762 TGCAGAGCTTCTTCAAGGCTTCTGGCAGTGGAGGCTGAAGACCTTACCTTGTATGATGA 821
QY 171 uGluSerSerLysArgProLeuGlyLeuLeuAlaArgGlnAlaGluAsnHisTyrAspG1 191
DB 822 GGAGTCTTCAAAAGGCCCTTCTGGCCCTTCTTGCAGACAAGCAGACCACTATGACCA 881
QY 191 nAspLeuAspGluLeuGlnLeuMetGluAspSerLysProHisProSerValGlnCy 211
DB 882 GGACCTGGATGAGCTCCAGCTGGAGATGGAGGACTCAAAGCCACACCCAGTGCCAGTG 941
QY 211 sSerProThrProGlyProPheLysProCysGluTyrLeuPheGluSerTrpGlyIleAr 231
DB 942 TAGCCCTACTCCAGGCCCTTCAAGCCCTGTGAGTACTCTTTGAAAGCTGGGGCATCCG 1001
QY 231 gLeuAlaValTrpAlaIleValLeuLeuSerValLeuCysAsnGlyLeuValLeuLeuTh 251
DB 1002 CTTGGCCGTGTGGGCCATCGTTGCTCTCGTCTCGCAATGCACTGGTGTCTGTGAC 1061
QY 251 rValPheAlaGlyGlyProAlaProLeuProProValLysPheValValGlyAlaIleAl 271
DB 1062 CGTGTTCGTGGCGGGGCTTGCCTCCCTGCCCCCGGTCAAGTTTGTGTAGTGGATGCGATGC 1121
QY 271 aGlyAlaAsnThrLeuThrGlyIleSerCysGlyLeuLeuAlaSerValAspAlaLeuTh 291
DB 1122 AGCGCCCAACACCTTGACTGGCATTTCTGTGGCTTCTAGCCTCAGTCGATGCCCTGAC 1181
QY 291 rPheGlyGlnPheSerGluTyrGlyAlaArgTrpGluThrGlyLeuGlyCysArgAlaTh 311
DB 1182 CTTTGGTCAGTTCTCTGAGTACGAGCCGCTGGAGACGGGGTAGGCTCCCTCCCTGGGCGAC 1241
QY 311 rGlyPheLeuAlaValLeuGlySerGluAlaSerValLeuLeuLeuThrLeuAlaIleVa 331
DB 1242 TGGCTTCTCGGCAGTACTTGGGTGGAGGCATCGGTGCTGCTGCTCACTCTGGCGCGAGT 1301
QY 331 lGlnCysSerValSerCysValArgAlaTyrGlyLysSerProSerLeuGlyLys 351
DB 1302 GCAGTGCAGCGCTCTCCGTCTCTGTGTCGGGCTATGGGAAGTCCCTCCCTGGGCGAG 1361
QY 351 rValArgAlaGlyValLeuGlyCysLeuAlaLeuAlaGlyLeuAlaAlaLeuProLe 371
DB 1362 CGTTCCAGCAGGGGTCTAGGCTGCTGCGCCTGGCAGTGGCAGGCTGGCGCGCCTGAGGG 1421
QY 371 uAlaSerValGlyGluTyrGlyAlaSerProLeuCysLeuProTyrAlaProProGluG1 391
DB 1422 GGCTCAGTGGGAAATACGGGCTCCCTCCCTGCTGCTGCTGCTGCTGCTGCTGCTGAGGG 1481
QY 391 YGlnProAlaAlaLeuGlyPheThrValAlaLeuValMetMetAsnSerPheCysPheLe 411
DB 1482 TCAGCCAGCAGCCCTGGGCTTCCCGTGGCCCTGGTGGTGGTGGTGGTGGTGGTGGTGGT 1541
QY 411 uValValAlaGlyAlaTyrIleLysLeuTyrCysAspLeuProArgGlyAspPheGluAl 431
DB 1542 GGTGTGGCGCGTGCCTTACATCAAACTGTAAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1601
QY 431 aValTrpAspCysAlaMetValArgHisValAlaTrpLeuIlePheAlaAspGlyLeuLe 451
DB 1602 CGTGTGGAGCTGGCGCATGGTGGGCGACGTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1661
QY 451 uTyrCysProValAlaPheLeuSerPheAlaSerMetLeuGlyLeuPheProValThrPr 471
DB 1662 CTACTGTCCGTGGCTTCTCAGCTTCCGCTTCCGCTTCCGCTTCCGCTTCCGCTTCCGCTTCC 1721
QY 471 oGluAlaValLysSerValLeuLeuValValLeuProLeuProAlaCysLeuAsnProLe 491
DB 1722 CGAGGCGGTCAAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1781
QY 491 uLeuTyrLeuLeuPheAsnProHisPheArgAspAspLeuArgLeuArgProArgAl 511
DB 1782 GCTGTACCTGCTCTTCAACCCCTTCCGCGGATGACCTTCCGCGGCTTCCGCGGCTTCCGCGG 1841
QY 511 aGlyAspSerGlyProLeuAlaTyrAlaAlaGlyGluLeuGluLysSerSerCysAs 531
```

```
Db 1842 AGGGAGCTCAGGCGCCCTAGCCTATGCTGCGCGCGGAGCTGGAGAGAGCTCTCTGTGA 1901
Qy 531 pSerThrGlnAlaLeuValAlaPheSerAspValAspLeuLeuLeuGluAlaSerGluAl 551
Db 1902 TTCTACCCAGGCGCTGGTAGCCTTCTCTGATGTGATCTCTATCTTCTGGAGCTTCTGAAGC 1961
Qy 551 aGlyArgProProGlyLeuGluThrTyTGlyPheProSerValThrLeuLeuSerCysG 571
Db 1962 TGGCGCGCGCCCTGGGCTGGAGACCTATGGCTTCCCTCAGTGACCTCATCTCTGTCA 2021
Qy 571 nGlnProGlyAlaProArgLeuGluGlySerHisCysValGluProGluGlyAanHisPh 591
Db 2022 GCAGCCAGGCGCGCCCGAGCTGGAGGCGAGCCATTTGTGTAGAGCCAGAGGGGACCAT 2081
Qy 591 eGlyAsnProGlnProSerMetAspGlyGluLeuLeuLeuArgAlaGluGlySerThrPr 611
Db 2082 TGGGAACCCCAACCTCCATGGATGGAGAACTGCTGTGAGGGCAGAGGGATCTACGCC 2141
Qy 611 oAlaGlyGlyLeuSerGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGly 631
Db 2142 AGCAGGTGGAGCTTTGTGAGGGGTGGCGGCTTTTCAGGCCCTCTGGCTTGGCTTTC 2201
Qy 631 rHisVal 633
Db 2202 ACACGTG 2208

RESULT 9
US-09-851-595-12
; Sequence 12, Application US/09851595
; Publication No. US20030166047A1
; GENERAL INFORMATION:
; APPLICANT: Gu, Wei
; TITLE OF INVENTION: NOVEL G-PROTEIN COUPLED RECEPTORS AND USES THEREFOR
; FILE REFERENCE: MNI-080CP
; CURRENT APPLICATION NUMBER: US/09/851,595
; CURRENT FILING DATE: 2000-05-08
; PRIOR APPLICATION NUMBER: 09/556,588
; PRIOR FILING DATE: 2000-05-08
; PRIOR APPLICATION NUMBER: 60/132,896
; PRIOR FILING DATE: 1999-05-06
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 12
; LENGTH: 2901
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)...(2901)
US-09-851-595-12

Alignment Scores:
Pred. No.: 4,54e-308 Length: 2901
Score: 3045.50 Matches: 602
Percent Similarity: 92.61% Conservat: 12
Best Local Similarity: 90.80% Mismatches: 19
Query Match: 92.09% Indels: 32
DB: 10 Gaps: 4

US-10-664-667-5 (1-633) x US-09-851-595-12 (1-2901)
Qy 1 AsnThrThrHisTyr-----5
Db 917 AACTCCACACACTATCTCTGAATGGTCCATGGACATCCAGGAGTTTCCAGATCTCAAG 976
Qy 6 ---ArgGluSerTrpTyAla-----CysArgTyrArg 15
Db 977 GCACACACAGCTGGAGATCTCTGACCTGACCCCGCGAGGATCCGGCTGCTCCCATCG 1036
Qy 16 -----SerGlyIleProGlySer--ThrHisAlaSerValGluArgSerGln---G 31
Db 1037 GGATGTGCCAACAGCTGCCCGAGGCTCCGAGTCTCGGAACACTGTCTCACAATCAAAATTGAGG 1096
```

```
Qy 31 yLeuSerLeuProAlaHisProAlaSerLeuAlaAlaLeuAlaAlaSerAsnThrThrAl 51
Db 1097 AGCTGCCCGAGCTGCACAGGTGTGAGAAATTTGGAGAAA--TCGGCTCCAAACACACCGC 1155
Qy 51 aSerGlyLeuGlu***AspThrPheSerGlnLeuSerSerLeuGlnAlaLeuAspLe 71
Db 1156 ATCTGGGAAATTTGAGCT--GACACCTTCAGCCAGCTGAGCTCCCTGCAAGCCCTCGATCT 1214
Qy 71 uSerTrpAsnAlaIleArgSerIleHisProGluAlaPheSerThrLeuHisSerLeuVa 91
Db 1215 TAGCTGGAAACCCATCCGTCCTCCATCCAGCTTCCAGGCTTCTCCACCTGCGACTCCCTGGT 1274
Qy 91 llysLeuAspLeuThrAspAsnGlnLeuThrThrLeuProLeuAlaGlyLeuGlyLe 111
Db 1275 CAAGCTGACCTGACAGACAAACAGCTGACACACTGCCCTGGCTGGCTGAGCTTGGGGCTT 1334
Qy 111 uMetHisLeuLeuLeuLeuGlyAsnLeuAlaLeuSerGlnAlaPheSerLeuAspSerPh 131
Db 1335 GATGATCTGAAGCTCAAAAGGGAACCTTGTCTCTCCAGGCTTCTCCAAAGGACAGTTT 1394
Qy 131 eProLysLeuArgIleLeuGluValProTyAlaTyrglnCysCysProTyrglnMetCy 151
Db 1395 CCCAAACTGAGGATCTGGAGGTGCCTTATGCTTACCAGTGTCTCTATGGATGTG 1454
Qy 151 sAlaSerPhePheLysAlaSerGlyGlnTrpGluAlaGluAspLeuHisLeuAspAspGl 171
Db 1455 TGCCAGCTTCTTCAAGGCTCTGGGAGTCTGGAGGCTGAAGACCTTCACTTGTATGATGA 1514
Qy 171 uGluSerSerLysArgProLeuGlyLeuLeuAlaArgGlnAlaGluAanHisTyrAspGl 191
Db 1515 GGAGTCTTTCAAAAAGGCGCCCTGGGCTCTTTCAGACAAACAGAGAAACCACTATGACCA 1574
Qy 191 nAspLeuAspGluLeuGlnLeuGluMetGluAspSerLysProHisProSerValGlnCy 211
Db 1575 GGACTGGATGAGCTCCAGCTGGAGATGGAGACTCAAAAGCCACACCCAGCTGCCAGTG 1634
Qy 211 sSerProThrProGlyProPheLysProCysGluTyrlLeuPheGluSerTrpGlyIleAr 231
Db 1635 TAGCCCTACTCCAGGCGCCCTTCAAGCCCTGTGAGTACCTCTTTGAAAGCTGGGCGATCCG 1694
Qy 231 gLeuAlaValTrpAlaIleValLeuLeuSerValLeuCysAsnGlyLeuValLeuLeuTh 251
Db 1695 CCTGGCGGTGGGCCCATCGTGTGCTCTCGCTCTGCAATGGAGCTGGTGTCTGCTGAC 1754
Qy 251 rValPheAlaGlyGlyProAlaProLeuProProValLysPheValValGlyAlaIleAl 271
Db 1755 CGTGTTCGTGGCGGGCTGCCCTCCCTGCCCGGTCAAGTTTGTGTAGTGGATGCGATTGC 1814
Qy 271 aGlyAlaAsnThrLeuThrGlyIleSerCysGlyLeuLeuAlaSerValAspAlaLeuTh 291
Db 1815 AGGCGCAACACCTTGACTGGCAITTTCTGTGGCTTCTAGCCCTCAGTCCGATGCCCTGAC 1874
Qy 291 rPheGlyGlnPheSerGluTyrglyAlaArgTrpGluThrGlyLeuGlyCysArgAlaTh 311
Db 1875 CTTTGGTCAGTTCTCTGAGTACGGAGCCCGCTGGAGAGCGGGGTAGGCTCCCGGGCCAC 1934
Qy 311 rGlyPheLeuAlaValLeuGlySerGluAlaSerValLeuLeuLeuThrLeuAlaAlaVa 331
Db 1935 TGGCTTCTGGCAGTACTTGGGTGGAGGCAATCGGTGCTGTCTGCTCACTCTGGCGCGAGT 1994
Qy 331 lGlnCysSerValSerValSerCysValArgAlaTyrglyLysSerProSerLeuGlySe 351
Db 1995 GCAGTGCAGCGCTCTCCGTCCTCTGTCCGGGCTATGGGAAGTCCCTCCCTCCCTGGGCGAG 2054
Qy 351 rValArgAlaGlyValLeuGlyCysLeuAlaLeuAlaGlyLeuAlaAlaLeuProLe 371
Db 2055 CGTTTCGAGCAGGGGTCTTAGGCTGCTGGCACTGGCAGGGGTGGCGCGCACTGCCCT 2114
Qy 371 uAlaSerValGlyGluTyrglyAlaSerProLeuCysLeuProTyrglyAlaProProGluGl 391
Db 2115 GGCCTCAGTGGGAAATACGGGGCTCTCCCACTCTGCTGCTGCTGCTGCTGCTGCTGCTGAGG 2174
Qy 391 yGlnProAlaAlaLeuGlyPheThrValAlaLeuValMetMetAsnSerPheCysPheLe 411
```



```
Db 2175 TCAGCCAGCCCTTGGCGCTTCCACCGTGGCCCTGGTGATGATGAACCTCTTCTGTTTCT 2234
Qy 411 uValValAlaGlyAlaTyrIleLeuLeuTyrCysAspLeuProArgGlyAspPheGluAl 431
Db 2235 GGTGTCGCGGTGCTACATCAAACTGTACTGTGACCTGCGCGGGCGACTTTGAGGC 2294
Qy 431 aValTyrAspCysAlaMetValArgHisValAlaTyrIleLeuPheAlaAspGlyLeuLe 451
Db 2295 CGTGTGGGACTGCGGCATGGTGAGCACGTGGCTCATCTTCGAGACGGGCTCT 2354
Qy 451 uTyrCysProValAlaPheLeuSerPheAlaSerMetLeuGlyLeuPheProValThrPr 471
Db 2355 CTACTGTCCCGTGGCTTCTCAGCTTGCCTCCATGCTGGCCCTTCTTCCCTGTCAAGCC 2414
Qy 471 oGluAlaValLysSerValLeuLeuValValLeuProLeuProAlaCysLeuAsnProLe 491
Db 2415 CGAGCGCTCAAGTCTGTCTGTGCTGTGGTGTGCTGCCCTGCTGCTGCCCTCAACCCACT 2474
Qy 491 uLeuTyrLeuLeuPheAsnProHisPheArgAspLeuArgLeuArgProAla 511
Db 2475 GCTGTACCTGCTTCTCAACCCCTTCCGGATGACCTTCGGCGGCTTCGGCCCGCGC 2534
Qy 511 aGlyAspSerGlyProLeuAlaTyrAlaAlaAlaGlyLeuLeuGlyLysSerSerCysAs 531
Db 2535 AGGGGACTCAGGGCCCTAGCTATGCTGCGCGGGGAGCTGGAGAAGACTCTCTGTGA 2594
Qy 531 pSerThrGlnAlaLeuValAlaPheSerAspValAspLeuIleLeuGluAlaSerGluAl 551
Db 2595 TTCTACCCAGGCCCTTGTAGGCTTCTCTGATGTGGATCTCATTTCTGGAAGCTTCTGAAGC 2654
Qy 551 aGlyArgProProGlyLeuGluThrTyrGlyPheProSerValThrIleLeuSerCysGl 571
Db 2655 TGGCGCGGCCCTGGCGTGGAGACTATGCGTCTCCCTCAGTGACCTCTCTCTGTCA 2714
Qy 571 nGlnProGlyAlaProArgLeuGluGlySerHisCysValGluProGluGlyAsnHisPh 591
Db 2715 GCAGCCAGGGCCCTCAGGCTGGAGGAGCCATTTGTAGAGCCAGAGGGGACCACTT 2774
Qy 591 eGlyAsnProGlnProSerMetAspGlyGlyLeuLeuLeuAlaGluGlySerThrPr 611
Db 2775 TGGGAACCCCAACCTCCATGGATGGAGAACTGCTGCTGAGGGCAGAGGGATCTAGGCC 2834
Qy 611 oAlaGlyGlyLeuSerGlyGlyGlyPheGlnProSerGlyLeuAlaPheAlaSe 631
Db 2835 AGCAGGTGGAGGCTTGTACAGGGGTGGCGGCTTTTCAGCCCTCTGGCTTGGCTTGTCTTC 2894
Qy 631 rHisVal 633
Db 2895 ACACGTG 2901
```

```
RESULT 10
US-10-664-667-12
; Sequence 12, Application US/10664667
; Publication No. US20040058377A1
; GENERAL INFORMATION:
; APPLICANT: Gu, wei
; FILE OF INVENTION: NOVEL G-PROTEIN COUPLED RECEPTORS AND USES THEREFOR
; FILE REFERENCE: NMI-080CP
; CURRENT APPLICATION NUMBER: US/10/664,667
; CURRENT FILING DATE: 2003-09-18
; PRIOR APPLICATION NUMBER: US/09/851,595
; PRIOR FILING DATE: 2000-05-08
; PRIOR APPLICATION NUMBER: 09/556,588
; PRIOR FILING DATE: 2000-05-08
; PRIOR APPLICATION NUMBER: 60/132,896
; PRIOR FILING DATE: 1999-05-06
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: PatentIn ver. 2.0
; SEQ ID NO 12
; LENGTH: 2901
; TYPE: DNA
; ORGANISM: Homo sapiens
```

```
;
; NAME/KEY: CDS
; LOCATION: (1)....(2901)
US-10-664-667-12

Alignment Scores:
Pred. No.: 4,54e-308 Length: 2901
Score: 3045.50 Matches: 602
Percent Similarity: 92.61% Conservative: 12
Best Local Similarity: 90.80% Mismatches: 19
Query Match: 92.09% Indels: 32
DB: 17 Gaps: 4

US-10-664-667-5 (1-633) x US-10-664-667-12 (1-2901)

Qy 1 AsnThrThrHisTyr----- 5
Db 917 AACTCCACACACTATCTCTGAATGGTGCCATGGACATCCAGGAGCTTTCCAGATCTCAAG 976
Qy 6 ---ArgGluSerTrpTyrAla----- CysArgTyrArg 15
Db 977 GCACCCACGCGCTGGAGATCCTGACCTGACCCCGCGCAGGCATCCGGCTGCTCCCATCGG 1036
Qy 16 -----SerGlyIleProGlySer-ThrHisAlaSerValGluArgSerGln---G1 31
Db 1037 GGATGTGCCAACAGCTGCCCGAGGCTCGAGTCTCGGAACCTGCTCACAATCAATTTGAGG 1096
Qy 31 YLeuSerLeuProAlaHisProAlaSerLeuAlaLeuAlaAlaSerAsnThrThrAl 51
Db 1097 AGCTGCCAGCCTGCACAGGTGTCAAGAAATGGAGGAAA-TCGGCCTTCCAACAACACCGC 1155
Qy 51 aSerGlyLysLeuGlu***AspThrPheSerGlnLeuSerSerLeuGlnAlaLeuAspLe 71
Db 1156 ATCTGGGAAATGGAGCT-GACACCTTCAGCCAGCTGAGCTCCCTGCAAGCCCTGGATCT 1214
Qy 71 uSerTrpAsnAlaIleArgSerIleHisProGluAlaPheSerThrLeuHisSerLeuVa 91
Db 1215 TAGCTGGAAAGCCATCCGTTCCATCCACCTTGAGGCTTCTCCACCTTGCACTCCCTGGT 1274
Qy 91 LysLeuAspLeuThrAspAsnGlnLeuThrThrLeuProLeuAlaGlyLeuGlyLe 111
Db 1275 CAAGCTGGACCTGACACACAACAGCTGACACACTGCCCTGGCTGGCTGGCTGGCTGGCT 1334
Qy 111 uMetHisLeuLysLeuLysGlyAsnLeuAlaLeuSerGlnAlaPheSerLysAspSerPh 131
Db 1335 GATGCATCTGAAGCTCAAGGGAAACCTTGCTCTCCAGGCCCTTCTCCAGGGACAGTTT 1394
Qy 131 eProLysLeuArgIleLeuGluValProTyrAlaTyrGlnCysCysProTyrGlyMetCy 151
Db 1395 CCCAAACTGAGGATCTGGAGGTGCTTATGCTACCACTGCTGCTGCTATGGATGTG 1454
Qy 151 sAlaSerPhePheLysAlaSerGlyGlnTrpGluAlaGluAspLeuHisLeuAspG1 171
Db 1455 TGGCAGCTTCTTCAAGGCCCTCTGGGAGTGGAGGCTGGAAGACCTTCACTTGTATGATGA 1514
Qy 171 uGluSerSerLysArgProLeuGlyLeuLeuAlaArgGlnAlaGluAsnHisTyrAspG1 191
Db 1515 GAGTCTTCAAAAAGGCCCTTGGGCTCTTGGCAGACAAGAGAGAACCACTATGACCA 1574
Qy 191 nAspLeuAspGluLeuGlnLeuGluMetGluAspSerLysProHisProSerValGlnCy 211
Db 1575 GGACCTGGATGAGCTCCAGCTGGAGATGGAGGACTCAAGGCCACACCCAGTGTCCAGTG 1634
Qy 211 sSerProThrProGlyProPheLysProCysGluTyrLeuPheGluSerTrpGlyIleAr 231
Db 1635 TAGCCCTACTCCAGGGCCCTTCAAGGCCCTGTGAGTACCTCTTTGAAAGCTGGGCACTCCG 1694
Qy 231 gLeuAlaValTrpAlaIleValLeuLeuSerValLeuCysAsnGlyLeuValLeuLeuTh 251
Db 1695 CTTGGCCGTGGGCCATCGTGTGCTCTCCGCTCTGCAATGGAGCTGTGTGCTGTGAC 1754
Qy 251 rValPheAlaGlyProAlaProLeuProValLysPheValValGlyAlaIleAl 271
```


Db 1755 CGTGTTCCTGCGGGGCTGCCCGCCCTGCCCGGCTCAAGTTTGGTAGGTGGCAATTGC 1814
Qy aGlyAlaAsnThrLeuThrGlyIleSerCysGlyLeuLeuAlaSerValAspAlaLeuTh 291
Db 1815 AGGCGCCAAACCTTGACTGGCAATTCCTGTGGCCTTCTAGCCTCAGTCGATGCCCTGAC 1874
Qy rPheGlyGlnPheSerGluThrGlyAlaArgTrpGluThrGlyLeuGlyCysArgAlaTh 311
Db 1875 CTTTGGTCAGTTCTCTAGTAGTCGAGCCGCTGGGAGACGGGGCTAGGCTGCGGGCCAC 1934
Qy rGlyPheLeuAlaValLeuGlySerGluAlaSerValLeuLeuLeuLeuThrLeuAlaVa 331
Db 1935 TGGCTTCTCGCAGTACTTGGGTGGAGGCATCGGTGCTGCTGCTCACTCTGGCGCGAGT 1994
Qy 331 lGlnCysSerValSerValSerCysValArgAlaTyrGlyLySerProSerLeuGlySe 351
Db 1995 GCAGTGCAGCGTCTCCGTCTCTGTGTCGGGGCTATGGGAAGTCCCTCCCTGGGCGAG 2054
Qy 351 rValArgAlaGlyValLeuGlyCysLeuAlaLeuAlaGlyLeuAlaAlaLeuProLe 371
Db 2055 CGTTTCGACAGGGGTCTTAGGTGCTGGCCTGGCAGGGCTGGCGCGCAGCTGCCCT 2114
Qy 371 uAlaSerValGlyGluThrGlyAlaSerProLeuCysLeuProTyrAlaProGluGl 391
Db 2115 GGCCTCAGTGGAGAAATACGGGGCTCCCCACTCTGCCTGCGCTACGGCCGACCTGAGGG 2174
Qy 391 yGlnProAlaAlaLeuGlyPheThrValAlaLeuValMetMetAenSerPheCysPheLe 411
Db 2175 TCAGCCAGCAGCCCTGGGCTTCACCGTGGCCTGGTGATGATGAATCTCTCTGTTCCT 2234
Qy 411 uValValAlaGlyAlaTyrIleLysLeuTyrCysAspLeuProArgGlyAspPheGluAl 431
Db 2235 GGTGTCGGCCGGTGCCTACATCAACTGTACTGTGACCTGCGCGGGCGGACTTTAGGC 2294
Qy 431 aValTrpAspCysAlaMetValArgHisValAlaTrpLeuIlePheAlaAspGlyLeuLe 451
Db 2295 CGTGTGGGACTGCGCCATGGTGAGCAGCTGGCTGCTCATCTTCGCAGACGGGCTCCT 2354
Qy 451 uTyrCysProValAlaPheLeuSerPheAlaSerMetLeuGlyLeuPheProValThrPr 471
Db 2355 CTACTGTCCCGTGGCTTCCTCAGCTTCGCTCCATGTGTGGCCTCTTCCCTGTTCACGCC 2414
Qy 471 oGluAlaValLysSerValLeuLeuValValLeuProLeuProAlaCysLeuAenProLe 491
Db 2415 CGAGGCCGTCAAGTCTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTCAACCCACT 2474
Qy 491 uLeuTyrLeuLeuPheAenProHisPheArgAspAspLeuArgLeuArgProArgAl 511
Db 2475 GCTGTACCTGCTCTTCAACCCCACTTCGGGATGACCTTCGGCGGCTTCGGCCCCCGGC 2534
Qy 511 aGlyAspSerGlyProLeuAlaTyrAlaAlaGlyGluLeuGlyLysSerSerCysAs 531
Db 2535 AGGGACTCAGGGCCCCCTAGCTATGCTTGGCGGGGAGCTGGAGAAGAGCTCTCTGTGA 2594
Qy 531 pSerThrGlnAlaLeuValAlaPheSerAspValAspLeuIleLeuGluAlaSerGluAl 551
Db 2595 TTCTACCCAGGCCCTGGTAGCTTCTCTGATGTGATCTCATCTCTGGAAGCTTCTGAAGC 2654
Qy 551 aGlyArgProProGlyLeuGluThrTyrGlyPheProSerValThrLeuIleSerCysGl 571
Db 2655 TGGCGGGCCCCCTGGGCTGGAGACCTATGGGCTTCCCTCAGTGACCTCATCTCTGTCA 2714
Qy 571 nGlnProGlyAlaProArgLeuGlySerHisCysValGluProGluGlyAenHisPh 591
Db 2715 GCAGCGAGGGCCCCAGGCTGGAGGGCAGGCATTTGTATAGGCAGAGGGGAACCACTT 2774
Qy 591 eGlyAenProGlnProSerMetAspGlyGluLeuLeuLeuArgAlaGluGlySerThrPr 611
Db 2775 TGGNAACCCCAACCTCCATGGATGGAGAACTGCTGTGAGGGCAGAGGGATCTACGCC 2834
Qy 611 oAlaGlyGlyGlyLeuSerGlyGlyGlyPheGlnProSerGlyLeuAlaPheAlaSe 631
Db 2835 AGCAGGTGGAGGCTTGTTCAGGGGGTGGGGCTTTTCAGGCCCTCTGGCTTGGGCTTTGCTTC 2894

Qy 631 rHisVal 633
Db 2895 ACACGTG 2901

RESULT 11

US-10-225-567A-580
; Sequence 580, Application US/10225567A
; Publication No. US20030113798A1

GENERAL INFORMATION:

; APPLICANT: LifeSpan Biosciences

; APPLICANT: Brown, Joseph P.

; APPLICANT: Burmer, Glenna C.

; APPLICANT: Roush, Christine L.

; TITLE OF INVENTION: ANTIGENIC PEPTIDES AND ANTIBODIES FOR G PROTEIN-COUPLED RECEPTORS

; FILE REFERENCE: 1920-4-4

; CURRENT APPLICATION NUMBER: US/10/225.567A

; CURRENT FILING DATE: 2001-12-19

; PRIOR APPLICATION NUMBER: 60/257,144

; PRIOR FILING DATE: 2000-12-19

; NUMBER OF SEQ ID NOS: 2292

; SOFTWARE: PatentIn version 3.1

; SEQ ID NO 580

; LENGTH: 3119

; TYPE: DNA

; ORGANISM: Homo sapiens

US-10-225-567A-580

Alignment Scores:

Pred. No.:	5.03e-308	Length:	3119
Score:	3045.50	Matches:	602
Percent Similarity:	92.61%	Conservative:	12
Best Local Similarity:	90.80%	Mismatches:	19
Query Match:	92.09%	Indels:	32
DB:	15	Gaps:	4

US-10-664-667-5 (1-633) x US-10-225-567A-580 (1-3119)

Qy 1 AsnThrThrHisTyr-----5

Db 802 AACTCCACACACTATCTGTAATGGTCCATCCAGGATTTCCAGATCTCAAG 861

Qy 6 ---ArgGluSerTrpTyrAla-----CysArgTyrArg 15

Db 862 GCACCCAGCCTGGAGATCTGACCCCTGACCCCGCGGAGCATCCGGTGTCTCCCATCGG 921

Qy 16 -----SerGlyIleProGlySer-ThrHisAlaSerValGluArgSerGln---G1 31

Db 922 GGATGTGCCAACAGCTGCCAGGCTCCGAGTCTCGGAACCTGTCTCACAATCAAAATTGAGG 981

Qy 31 yLeuSerLeuProAlaHisProAlaSerLeuAlaAlaLeuAlaAlaSerAsnThrThrAl 51

Db 982 AGTGTCCAGCCTGCACAGGTGTGAGAAATTGGAGGAAA--TCGGCCTCCAACACACCGC 1040

Qy 51 aSerGlyLysLeuGlu***AspThrPheSerGlnLeuSerSerLeuGlnAlaLeuAspLe 71

Db 1041 ATCTGGAAATTGGAGCT-GACACCTTCAGCCAGCTGAGCTCCCTGCAAGCCCTGGAATCT 1099

Qy 71 uSerTrpAsnAlaIleArgSerIleHisProGluAlaPheSerThrLeuHisSerLeuVa 91

Db 1100 TAGCTGGAAACCCATCCGGTCCATCCACCTTGAGGCCCTTCTCCACCTGCACTCCCTGCT 1159

Qy 91 lLysLeuAspLeuThrAspAenGlnLeuThrLeuProLeuAlaGlyLeuGlyLe 111

Db 1160 CAAGCTGACCTGACAGACAAACAGCTGACCACTGCCCCCTGGCTGGACTTGGGGGCTT 1219

Qy 111 uMetHisLeuLysLeuLysGlyAsnLeuAlaLeuSerGlnAlaPheSerLysAspSerPh 131

Db 1220 GATCATCTGAAGCTCAAGGGAACCTTGTCTCTCCAGGCCCTTCTCCAGGACAGTTT 1279

Qy 131 eProLysLeuArgIleLeuGluValProTyrAlaTyrGlnCysCysProTyrGlyMetCy 151

Db 1280 CCCAAAACCTGAGGATCTGGAGGTGCCTTATGCTTACCAGTGTCTTCCCTATGGGATGTG 1339

151 sAlaserPhePheLysAlaserGlyGlnTrpGluAlaGluAspLeuHisLeuAspAspG1 171
 1340 TGCAGGCTTCTCAAGGCTCTGGGAGCTGGAGGCTGAAGACCTTACCTTGATGATGA 1399
 171 uGluSerSerLysArgProLeuGlyLeuLeuAlaArgGlnAlaGluAsnHisTyrAspG1 191
 1400 GGAGTCTTCAAAAAGGCCCTTGGGCTCTTGGCCAGACAAGCAGAGAACCACCTATGACCA 1459
 191 nAspLeuAspGlnLeuGlnLeuGluMetGluAspSerLysProHisProSerValGlnCy 211
 1460 GGACCTGATGAGCTCCAGCTGGAGATGGAGACTCAAAGCCACACCCAGTGTCCAGTG 1519
 211 sSerProThrProGlyProPheLysProCysGluTyrLeuPheGluSerTrpGlyVileAr 231
 1520 TAGCCCTACTCCAGGCCCTTCAAGCCCTGTGAGTACTCTTTGAAGACTGGGGCATCCG 1579
 231 gLeuAlaValTrpAlaIleValLeuLeuSerValLeuCysAsnGlyLeuValLeuLeuTh 251
 1580 CCTGGCCGTGTGGGCCATCGTGTGCTCTCGTGTCTCTGCAATGACTGGTGTGCTGTGAC 1639
 251 rValPheAlaGlyGlyProAlaProLeuProProValLysPheValValGlyAlaIleAl 271
 1640 CGTGTTCGTGGCGGGCTGCCCCCTGCCCCCGTCAAGTGTGTGGTAGGTGGGATTGC 1699
 271 aGlyAlaAsnThrLeuThrGlyIleSerCysGlyLeuLeuAlaSerValAspAlaLeuTh 291
 1700 AGGCCCAACACCTTGACTGGCATTTCTGTGGCTCTTAGCCCTCAGTCGATGCCCTGTGAC 1759
 291 rPheGlyGlnPheSerGluTyrGlyAlaArgTrpGluThrGlyLeuGlyCysArgAlaTh 311
 1760 CTTTGGTCAGTTCTCTGAGTACGGAGCCGCTGGAGACGGGGCTAGGCTGCCGGGCCAC 1819
 311 rGlyPheLeuAlaValLeuGlySerGluAlaSerValLeuLeuLeuThrLeuAlaIaVa 331
 1820 TGGCTTCTCGCAGTACTTGGGTCGAGGATCGGTCTGTCTCCTCTGGCGCGCAGT 1879
 331 lGlnCysSerValSerCysValArgAlaTyrGlyLysSerProSerLeuGlySe 351
 1880 GCAGTGACGGCTCTCCGTCTCTGTGTCGGGCTATGGGAAGTCCCCCTCCCTGGGCGAG 1939
 351 rValArgAlaGlyValLeuGlyCysLeuAlaLeuAlaGlyLeuAlaAlaLeuProLe 371
 1940 CGTTCGAGCAGGGCTCTAGGCTGCTGGCACTGGCAGGGCTGGCCGCGCACTGCCCT 1999
 371 uAlaSerValGlyGluTyrGlyAlaSerProLeuCysLeuProTyrAlaProProGluG1 391
 2000 GGCTCTAGTGGGAGATACGGGGCTCCCACTCTGCCCTGACCTGACGCGCACCTGAGGG 2059
 391 yGlnProAlaAlaLeuGlyPheThrValAlaLeuValMetMetAsnSerPheCysPheLe 411
 2060 TCAGCCAGACGCTTGGCTTACCGTGGCTCTGGTATGATGAACCTCTCTGTTCCT 2119
 411 uValValaGlyAlaTyrIleLysLeuTyrCysAspLeuProArgGlyAspPheGluAl 431
 2120 GGTGTCGGCGGTCCTTACATCAACTGTACTGTGACCTGCCGGGGCGGACTTTGAGGC 2179
 431 aValTrpAspCysAlaMetValArgHisValAlaTrpLeuPheAlaAspGlyLeuLe 451
 2180 CGTGTGGGACTTGGCCATGTGTAGGCACGTGGCTCTCATCTTCGACAGCGGGCTCCT 2239
 451 uTyrCysProValAlaPheLeuSerPheAlaSerMetLeuGlyLeuPheProValThrPr 471
 2240 CTACTGTCCCGTGGCTTCTCTACGCTTGGCTCTCATGCTGGGCTCTTCCCTGTGACGCC 2299
 471 oGluAlaValLysSerValLeuLeuValLeuProLeuProAlaCysLeuAsnProLe 491
 2300 CGAGCCGCTCAAGTCTGTCTGTGGTGGTGTGCTGCTGCTGCTGCTGCTCAACCCACT 2359
 491 uLeuTyrLeuLeuPheAsnProHisPheArgAspAspLeuArgLeuArgProArgAl 511
 2360 GCTGTACTCTGCTTCAACCCCACTTCCGGAGTACCTTCCGGCGGCTTCCGGCCCCCGC 2419

511 aGlyAspSerGlyProLeuAlaTyrAlaAlaAlaGlyGluLeuGlyLysSerSerCysAs 531
 2420 AGGGGACTCAGGGCCCTTAGCCTATGCTGGCGCGGGAGCTGGAGAAGAGCTCTCTGTA 2479
 531 pSerThrGlnAlaLeuValAlaPheSerAspValAspLeuLeuGluAlaSerGluAl 551
 2480 TTCTACCCAGGCCCTGTGTAGCTTCTCTGATGTGATCTCATTTCTGGAAGCTTCTGAAGC 2539
 551 aGlyArgProGlyLeuGluThrTyrGlyPheProSerValThrLeuIleSerCysG1 571
 2540 TGGCGCGCCCTTGGGCTGGAGACCTATGGCTTCCCTCAGTGACCTCATCTCTCTGTCA 2599
 571 nGlnProGlyValaProArgLeuGluGlySerHisCysValGluProGluGlyAsnHisPh 591
 2600 GCAGCCAGGGGCCCTCAGGCTGGAGGCGAGCCATTGTGTAGAGCCAGAGGGGACCACTT 2659
 591 eGlyAsnProGlnProSerMetAspGlyGluLeuLeuLeuLeuArgAlaGluGlySerThrPr 611
 2660 TGGGAACCCCAACCTTCCATGATGGAGAACTGCTGCTGAGGGCAGAGGATCTACGCC 2719
 611 oAlaGlyGlyLeuSerGlyGlyGlyPheGlnProSerSerCylLeuAlaPheAlaSe 631
 2720 AGCAGGTGGAGGCTTGTTCAGGGGGTGGCGCTTTCAGCCCTCTGGCTTGGCTTGGCTTC 2779
 631 rHisVal 633
 2780 ACAGTG 2786
 RESULT 12
 US-10-331-496A-52
 ; Sequence 52, Application US/10331496A
 ; Publication No. US20030228305A1
 ; GENERAL INFORMATION:
 ; APPLICANT: FRANTZ, GRETCHEN
 ; APPLICANT: HILLAN, KENNETH J.
 ; APPLICANT: PHILLIPS, HEIDI S.
 ; APPLICANT: POLAKIS, PAUL
 ; APPLICANT: SMITH, VICTORIA
 ; APPLICANT: SPENCER, SUSAN D.
 ; APPLICANT: WILLIAMS, P. MICKEY
 ; APPLICANT: WU, THOMAS D.
 ; APPLICANT: ZHANG, ZEMIN
 ; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE DIAGNOSIS AND
 ; FILE REFERENCE: P5014R1-PCT
 ; CURRENT APPLICATION NUMBER: US/10/331,496A
 ; CURRENT FILING DATE: 2002-12-30
 ; PRIOR APPLICATION NUMBER: US 60/345,444
 ; PRIOR FILING DATE: 2002-01-02
 ; PRIOR APPLICATION NUMBER: US 60/351,885
 ; PRIOR FILING DATE: 2002-01-25
 ; PRIOR APPLICATION NUMBER: US 60/360,066
 ; PRIOR FILING DATE: 2002-02-25
 ; PRIOR APPLICATION NUMBER: US 60/362,004
 ; PRIOR FILING DATE: 2002-03-05
 ; PRIOR APPLICATION NUMBER: US 60/366,869
 ; PRIOR FILING DATE: 2002-03-20
 ; PRIOR APPLICATION NUMBER: US 60/366,284
 ; PRIOR FILING DATE: 2002-03-21
 ; PRIOR APPLICATION NUMBER: US 60/368,679
 ; PRIOR FILING DATE: 2002-03-28
 ; PRIOR APPLICATION NUMBER: US 60/404,809
 ; PRIOR FILING DATE: 2002-08-19
 ; PRIOR APPLICATION NUMBER: US 60/405,645
 ; PRIOR FILING DATE: 2002-08-21
 ; NUMBER OF SEQ ID NOS: 95
 ; SEQ ID NO 52
 ; LENGTH: 3325
 ; TYPE: DNA
 ; ORGANISM: Homo sapien
 US-10-331-496A-52
 Alignment Scores:

Pred. No.: 5.51e-308 Length: 3325
Score: 3045.50 Matches: 602
Percent Similarity: 92.61% Conservative: 12
Best Local Similarity: 90.80% Mismatches: 19
Query Match: 92.05% Indels: 32
DB: 17 Gaps: 4

US-10-664-667-5 (1-633) x US-10-331-496A-52 (1-3325)

QY 1 AenThrThrHisTyr-----5
DB 845 AACTCCACACATATCTCTGATGTGCGCATCCAGGAGTTTCCAGATCTCAAG 904
QY 6 ---ArgGluSerTyrAla-----CysArgTyrArg 15
DB 905 GCACACCAGCTCGAGATCTGACCTGACCCGCGCAGGATCCGGCTGCCATCGG 964
QY 16 -----SerGlyIleProGlySer-ThrHisAlaSerValGluArgSerGln---G1 31
DB 965 GGATGTGCCAACAGCTGCCAGGCTCCGAGTCTCGAATCTGCTCACAATCAAAATTGAGG 1024
QY 31 YLeuSerLeuProAlaHisProAlaSerLeuAlaAlaLeuAlaAlaSerAsnThrAl 51
DB 1025 AGCTGCCAGCTGCACAGGTGTGAGAAATGGAGGAAA-TCGGGCTCCACACACACCGC 1083
QY 51 aSerGlyLeuGlu***AspThrPheSerGlnLeuSerSerLeuGlnAlaLeuAspLe 71
DB 1084 ATCTGGGAAATTGGAGCT-GACACCTTCAGCCAGCTGAGCTCCCTGCAAGCCCTGGATCT 1142
QY 71 uSerTrpAsnAlaIleArgSerIleHisProGluAlaPheSerThrLeuHisSerLeuVa 91
DB 1143 TAGCTGGAAACGCCATCCGGTCCATCCACCTGAGGCTTCTCCACCTGCACTCCCTGGT 1202
QY 91 lLeuLeuAspLeuThrAspAsnGlnLeuThrThrLeuProLeuAlaGlyLeuGlyLe 111
DB 1203 CAAGCTGGACCTGACAGAACACAGCTGACACACTGCCCTGGCTGGCTGGGGGCTT 1262
QY 111 uMetHisLeuLeuLeuGlyAsnLeuAlaLeuSerGlnAlaPheSerLeuAspSerPh 131
DB 1263 GATGCATCTGAAGCTCAAGGGAACCTTGTCTCTCCAGGCTTCTCCAGGACAGTTT 1322
QY 131 eProlyLeuArgIleLeuGluValProTyrAlaTyrGlnCysCysProTyrGlyMetCy 151
DB 1323 CCCAAAAGTGGAGTCTCGAGGTGCCTTATGCTTACCAGTGTCTGCTTATGGGATGTG 1382
QY 151 eAlaSerPheLeuAlaSerGlyGlnTrpGluAlaGluAspLeuHisLeuAspAspG1 171
DB 1393 TGCCAGCTTCTCAAGGCTCTGGGAGTGGAGGCTGAAGACCTTCACCTTTGATGATGA 1442
QY 171 uGluSerSerLeuArgProLeuGlyLeuAlaArgGlnAlaGluAsnHisTyrAspG1 191
DB 1443 GGAGTCTTCAAAAAGGCCCTGGGCTCTTGGCCAGACAGACAGAACCACTATGACCA 1502
QY 191 nAspLeuAspGluLeuGlnLeuGluMetGluAspSerLeuProHisProSerValGlnCy 211
DB 1503 GGACCTGTAGCTTCCAGCTCGAGTGGAGTGGAGGACTCAAAAGCCACACCCAGTGTCCAGTG 1562
QY 211 eSerProThrProGlyProPheLeuProCysGluTyrLeuPheGluSerTrpGlyIleAr 231
DB 1563 TAGCCCTACTCCAGGCCCTTCAAGCCCTTGAGTACCTCTTTGAAAGCTGGGGATCCG 1622
QY 231 gLeuAlaValTrpAlaIleValLeuLeuSerValLeuCysAsnGlyLeuValLeuLeuTh 251
DB 1623 CTGGCCGTGGGCCCATCGTGTCTCTCGGTGCTCTGCAATGGACTGGTGTGCTGAC 1682
QY 251 rValPheAlaGlyGlyProAlaProLeuProProVallylsPheValValAlaIleAl 271
DB 1683 CGTGTTCGCTGGCGGCTGCCCCCTGCCCGGCTCAAGTTTGTGGTAGGTGCGATTGC 1742
QY 271 aGlyAlaAsnThrLeuThrGlyIleSerCysGlyLeuLeuAlaSerValAspAlaLeuTh 291
DB 1743 AGGCGCCAAACCTTTGACTGGCATTTCTGTGGCTTCTAGCCTCAGTCGATGCCCTGAC 1802

RESULT 13

QY 291 rPheGlyGlnPheSerGluTyrGlyAlaArgTrpGluThrGlyLeuGlyCysArgAlaTh 311
DB 1803 CTTTGGTCAGTTCTCTAGTACGAGGCCGCTGGGAGACGGGGCTAGGCTGCCGGGCCAC 1862
QY 311 rGlyPheLeuAlaValLeuGlySerGluAlaSerValLeuLeuLeuThrLeuAlaAlaVa 331
DB 1863 TGGCTTCTTGGCAGTACTTGGGTGGAGGCATCGGTGCTGCTGCTCACTCTGGCCGAGT 1922
QY 331 lGlnCysSerValSerCysValArgAlaTyrGlyLySerProSerLeuGlySe 351
DB 1923 GCAGTGCAGCGCTCTCCGTCTCTGTGTCGGGCTATGGGAAGTCCCTCTCCCTGGGCG 1982
QY 351 rValArgAlaGlyValLeuGlyCysLeuAlaLeuAlaGlyLeuAlaAlaLeuProLe 371
DB 1983 CGTTTCGAGCAGGGTCTCTAGGCTGCTGGCACTGGCAGGGCTGGCGCGCAGCTGCCCT 2042
QY 371 uAlaSerValGlyGluTyrGlyAlaSerProLeuCysLeuProTyrAlaProGluG1 391
DB 2043 GGCTCTCAGTGGGAGAAATACGGGGCTCCCTCACTCTGCTGCTGCTGCTGCTGCTGAGG 2102
QY 391 yGlnProAlaAlaLeuGlyPheThrValAlaLeuValMetMetAsnSerPheCysPheLe 411
DB 2103 TCAGCCAGCAGCCCTGGGCTTCACCGTGGCTTGTGTGATGATGAATCTCTTGTTCCT 2162
QY 411 uValAlaAlaGlyAlaTyrIleLeuLeuTyrCysAspLeuProArgGlyAspPheGluAl 431
DB 2163 GGTGTCGGCGCTGCTACATCAAACTGTACTGTGACCTGCGCGGGGCGAGCTTTAGGC 2222
QY 431 aValTrpAspCysAlaMetValArgHisValAlaTrpLeuIlePheAlaAspGlyLeuLe 451
DB 2223 CGTGTGGAGCTGCGCCATGGTGGAGCAGTGGCTGCTGCTCATCTTCGAGACGGGCTCT 2282
QY 451 uTyrCysProValAlaPheLeuSerPheAlaSerMetLeuGlyLeuPheProValThrPr 471
DB 2283 CTACTGTCCCTGGGCTTCTCAGCTTCGCTTCCATGCTGGGCTCTTCTCTGTCACGCC 2342
QY 471 oGluAlaVallySerValLeuLeuValLeuProLeuProAlaCysLeuAsnProLe 491
DB 2343 CGAGCGCTCAAGTCTGTCTGCTGGTGTGCTGCTGCTGCTGCTGCTGCTCAACCCACT 2402
QY 491 uLeuTyrLeuLeuPheAsnProHisPheArgAspLeuArgLeuArgProArgAl 511
DB 2403 GCTGTACTGTCTTCAACCCCTTCCGGGATGACCTTCGGCGGCTTCGGCCCCGGCG 2462
QY 511 aGlyAspSerGlyProLeuAlaTyrAlaAlaGlyGluLeuGlyLySerSerCysAs 531
DB 2463 AGGGGACTCAGGGGCCCTAGCTATGCTGCGCGCGGGAGCTGGAGAGAGACTCTCTGTA 2522
QY 531 pSerThrGlnAlaLeuValAlaPheSerAspValAspLeuIleLeuGluAlaSerGluAl 551
DB 2523 TTCTACCCAGGCCCTGGTAGCTTCTCTGATGTGGATCTCATTTCTGGAAGCTTCTGAAGC 2582
QY 551 aGlyArgProGlyLeuGluThrThrGlyPheProSerValThrLeuIleSerCysG1 571
DB 2583 TGGCGCGCCCTCGGCTGGAGCTATGGCTTCCCTCAGTGACCTCTCATCTCTGTCA 2642
QY 571 nGlnProGlyAlaProArgLeuGlyLySerHisCysValGluProGluGlyAsnHisPh 591
DB 2643 GCAGCCAGGGGCCCTCAGGCTGGAGGGCAGCCATTTGTAGAGCCAGGGGGAACCACTT 2702
QY 591 eGlyAsnProGlnProSerMetAspGlyGluLeuLeuLeuArgAlaGluGlySerThrPr 611
DB 2703 TGGGAAACCCCAACCTCCATGGATGGAGAACTGCTGCTGAGGGCAGAGGGATCTAGGCC 2762
QY 611 oAlaGlyGlyLeuSerGlyGlyGlyPheGlnProSerGlyLeuAlaPheAlaSe 631
DB 2763 AGCAGGTGGAGGCTTGTGAGGGGTGGGGCTTTTCAAGCCCTCTGGCTTGGCTTGGCTTC 2822
QY 631 rHisVal 633
DB 2823 ACAGTG 2829

```
US-10-398-036-26
; Sequence 26, Application US/10398036
; Publication No. US20040137564A1
; GENERAL INFORMATION:
; APPLICANT: INCYTE GENOMICS INC.; BAUGHN, Mariah R.;
; APPLICANT: GRAUL Richard C.; CHAWLA, Narinder K.;
; APPLICANT: GANDHI, Ameena R.; HAFALIA, April J.A.;
; APPLICANT: RAMKUMAR, Javalaxmi; TRIBOULEY, Catherine M.;
; APPLICANT: THORNTON, Michael B.; KALLICK, Deborah A.;
; APPLICANT: YAO, Monique G.; ELLIOTT, Vicki S.;
; APPLICANT: BURFORD, Neil; KHAN, Farrah A.;
; APPLICANT: YUE, Henry; LU, Yan;
; APPLICANT: ARVIZU, Chandra S.; ROOPA, Reddy M.;
; APPLICANT: NGUYEN, Daniel B.; LEE, Ernestine A.;
; APPLICANT: LU, Dyung Aina M.; ISON, Craig H.;
; APPLICANT: WALSH, Roderick T.; POLICKY, Jennifer L.
; TITLE OF INVENTION: G-PROTEIN COUPLED RECEPTORS
; FILE REFERENCE: PI-0236 USN
; CURRENT APPLICATION NUMBER: US/10/398,036
; PENDING FILING DATE: 2003-03-28
; PRIOR APPLICATION NUMBER: PCT/US01/30661
; PRIOR FILING DATE: 2001-09-28
; PRIOR APPLICATION NUMBER: US 60/245,855
; PRIOR FILING DATE: 2000-11-03
; PRIOR APPLICATION NUMBER: US 60/242,322
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: US 60/240,589
; PRIOR FILING DATE: 2000-10-13
; PRIOR APPLICATION NUMBER: US 60/249,343
; PRIOR FILING DATE: 2000-11-15
; PRIOR APPLICATION NUMBER: US 60/247,587
; PRIOR FILING DATE: 2000-11-09
; PRIOR APPLICATION NUMBER: US 60/245,900
; PRIOR FILING DATE: 2000-11-03
; PRIOR APPLICATION NUMBER: US 60/242,223
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: US 60/236,546
; PRIOR FILING DATE: 2000-09-29
; NUMBER OF SEQ ID NOS: 32
; SOFTWARE: PERL Program
; SEQ ID NO 26
; LENGTH: 3429
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc.feature
; OTHER INFORMATION: Incyte ID No: 7479890CB1
US-10-398-036-26

Alignment Scores:
Pred. No.: 5,76e-308 Length: 3429
Score: 3045.50 Matches: 602
Percent Similarity: 92.61% Conservative: 12
Best Local Similarity: 90.80% Mismatches: 19
Query Match: 92.09% Indels: 32
DB: 18 Gaps: 4

US-10-664-667-5 (1-633) x US-10-398-036-26 (1-3429)

Qy 1 AsnThrThrHisTyr-----5
Db 860 AACTCCACACACTATCTGTAATGGTGCCATGCCAGGAGTTTCCAGGTCTCAAG 919

Qy 6 ---ArgGluSerTrpTyrAla-----CysArgTyrArg 15
Db 920 GCACCACAGCCTGGAGATCTGACCTGACCCGCGGAGGCATCCGGCTGCTCCCATCG 979

Qy 16 -----SerGlyIleProGlySer-ThrHisAlaSerValGluArgSerGln---G1 31
Db 980 GGATGTGCCAACAGCTGCCAGGCTCCGAGTCTCGGAACATGTCTCAACAATCAATGAGG 1039

Qy 31 YLeuSerLeuProAlaHisProAlaSerLeuAlaAlaAlaAlaSerAsnThrThrAl 51
Db 2118 TCAGCCAGCAGCCCTGGGCTTCACCGTGGCCCTGGTGATGATGAACCTCTCTTCTTCTTCT 2177
```

```
QY 411 uValValAlaGlyAlaTyrIleLysLeuTyrCysAspLeuProArgGlyAspPheGluAl 431
Db 2178 GGTCTGGCGGTGCTACATCAAACTGTACTGTGACCTGCGCGGGGCGACTTTGAGGC 2237
QY 431 aValTrpAspCysAlaMetValArgHisValAlaTrpLeuIlePheAlaAspGlyLeuLe 451
Db 2238 COTGTGGGACTGCGGCATGGTGGAGCAGTGGCTGCTCATCTTCGCAGACGGGCTCCT 2297
QY 451 uTyrCysProValAlaPheLeuSerPheAlaSerMetLeuGlyLeuPheProValThrPr 471
Db 2298 CTACTGTCCGTGGGCTTCTCAGCTTGGCTTCATGTCTGGGCTTCTTCCCTGTACGCC 2357
QY 471 oGluAlaValLysSerValLeuValLeuValLeuProLeuProAlaCysLeuAsnProLe 491
Db 2358 CGAGGCCGTCAAGTCTGTCTGTGGTGTGTGGTGTGGTGTGGTGTGGTGTGGTGTGGTGT 2417
QY 491 uLeuTyrIleLeuPheAsnProHisPheArgAspLeuArgLeuArgProArgAl 511
Db 2418 GCTGTACCTGTCTTCAACCCCTTCCGGGATGACCTTCGGCGGCTTCGGCCCGCGC 2477
QY 511 aGlyAspSerGlyProLeuAlaTyrAlaAlaGlyGluLeuLysSerSerCysAs 531
Db 2478 AGGGGACTCAGGGCCCTTAGCTATGCTGGCGCGGGAGCTGGAGAGAGAGCTTCTGTGA 2537
QY 531 pSerThrGlnAlaLeuValAlaPheSerAspValAspLeuIleLeuGluAlaSerGluAl 551
Db 2538 TTCTACCCAGGCCCTGGTAGCTTCTGTGTGGATCTCATCTTGTGAAGCTTCTGAAGC 2597
QY 551 aGlyArgProProGlyLeuGluThrTyrGlyPheProSerValThrLeuIleSerCysGl 571
Db 2598 TGGCGGGCCCTTGGGCTGGAGACTATGGCTTCCCTCAGTGACCTCATCTCTCTGTCA 2657
QY 571 nGlnProGlyAlaProArgLeuGluGlySerHisCysValGluProGluGlyAsnHisPh 591
Db 2658 GCAGCCAGGGGCCCCAGGCTGGAGGAGCCATTTGTAGAGCCAGAGGGGAGACCATTT 2717
QY 591 eGlyAsnProGlnProSerMetAspGlyGluLeuLeuLeuArgAlaGluGlySerThrPr 611
Db 2718 TGGGAACCCCAACCTCCATGGATGGAGAACTGCTGCTGAGGGCAGAGGGATCTACGCC 2777
QY 611 oAlaGlyGlyGlyLeuSerGlyGlyGlyPheGlnProSerGlyLeuAlaPheAlase 631
Db 2778 AGCAGGTGGAGCTTGTGAGGGGTGGGGCTTTTCAGGCCCTTGGCTTGGCTTGGCTTC 2837
QY 631 rHisVal 633
Db 2838 ACACGTG 2844
```

RESULT 14

```
US-10-176-847-89
; Sequence 89, Application US/10176847
; Publication No. US20030068636A1
; GENERAL INFORMATION:
; APPLICANT: Veiby, Petteur Ole
; TITLE OF INVENTION: COMPOSITIONS, KITS, AND METHODS FOR
; TITLE OF INVENTION: IDENTIFICATION, ASSESSMENT, PREVENTION, AND THERAPY OF BREAST
; TITLE OF INVENTION: AND OVARIAN CANCER
; FILE REFERENCE: MRI-039
; CURRENT APPLICATION NUMBER: US/10/176,847
; CURRENT FILING DATE: 2002-06-21
; NUMBER OF SEQ ID NOS: 112
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 89
; LENGTH: 3438
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-176-847-89

Alignment Scores:
Pred. No.: 5,78e-308 Length: 3438
Score: 3045.50 Matches: 602
Percent Similarity: 92.61% Conservative: 12
```

```
Best Local Similarity: 90.80% Mismatches: 19
Query Match: 92.09% Indels: 32
DB: 14 Gaps: 4
US-10-664-667-5 (1-633) x US-10-176-847-89 (1-3438)
QY 1 AsnThrThrHisTyr----- 5
Db 969 AACTCCACACACTACTCTGAATGGTCCATCGGACATCCAGGAGTTTCCAGATCTCAAG 1028
QY 6 ---ArgGluSerTrpTyrAla-----CysArgTyrArg 15
Db 1029 GCACCACAGCTCGAGATCTGACCTCGACCCCGGAGGAGCATCGGCTGTCTCCATCGG 1089
QY 16 -----SerGlyIleProGlySer-ThrHisAlaSerValGluArgSerGln---G1 31
Db 1089 GGATGTGCCAACAGCTGCCAGGCTCCGAGTCTCGGAACCTGTCTCACAATCAAATTGAGG 1148
QY 31 YLeuSerLeuProAlaHisProAlaSerLeuAlaAlaLeuAlaAlaSerAsnThrThrAl 51
Db 1149 AGCTGCCAGCTGCACAGGTGTAGAAATTGGAGGAAA--TCGGCCTCCAAACACACCGC 1207
QY 51 aSerGlyIleLeuGlu---AspThrPheSerGlnLeuSerSerLeuGlnAlaLeuAspLe 71
Db 1208 ATCTGGGAAATTGGAGCT-GACACCTTCAGCCAGCTGAGCTCCCTGCAAGCCCTGGATCT 1266
QY 71 uSerTrpAsnAlaIleArgSerIleHisProGluAlaPheSerThrLeuHisSerLeuVa 91
Db 1267 TAGCTGGNACCCATCCGTCATCCACCTCGAGGCCCTTCTCCACCTGCACTCCCTGGT 1326
QY 91 lLysLeuAspLeuThrAspAsnGlnLeuThrThrLeuProLeuAlaGlyLeuGlyLe 111
Db 1327 CAAGCTGACCTGCACAGAACACCACTGCACACACTGCCCTGGCTGAGCTTGGGGCTT 1386
QY 111 uMetHisLeuLysLeuGlyAsnLeuAlaSerGlnAlaPheSerLysAspSerPh 131
Db 1387 GATGCATCTGAAGCTCAAAAGGGAACCTTGTCTCTCCAGGCCCTTCTCCAAAGGACAGTTT 1446
QY 131 eProLysLeuArgIleLeuGluValProTyrAlaTyrGlnCysCysProTyrGlyMetCy 151
Db 1447 CCCAAACTGAGGATCCTGGAGGTGCCTTATGCTTACCAGTGTCTCCCTATGGATGTG 1506
QY 151 sAlaSerPhePheLysAlaSerGlyGlnTrpGluAlaGluAspLeuHisLeuAspAspGl 171
Db 1507 TGCCAGCTTCTTCAAGGCTCTGGGCGAGTGGAGGCTGAAGACCTTCACCTTGATGATGA 1566
QY 171 uGluSerSerLysArgProLeuGlyLeuLeuAlaArgGlnAlaGluAsnHisTyrAspGl 191
Db 1567 GGAGTCTTTCAAAAAGGCCCTCTGGGCCCTCTTGGCCAGACAAAGCAGAGAACCACTATGACCA 1626
QY 191 nAspLeuAspGluLeuGlnLeuGluMetGluAspSerLysProHisProSerValGlnCy 211
Db 1627 GGACTGGATGAGCTCCAGCTGGAGATGGAGACTCAAAGCCACACCCCACTGTCCAGTG 1686
QY 211 sSerProThrProGlyProPheLysProCysGluTyrLeuPheGluSerTrpGlyIleAr 231
Db 1687 TAGCCCTACTCCAGGCCCTTCAAGCCCTGTGAGTACCTCTTTGAAAGCTGGGGCATCCG 1746
QY 231 gLeuAlaValTrpAlaIleValLeuLeuSerValLeuCysAsnGlyLeuValLeuLeuTh 251
Db 1747 CCTGGCGGTGTGGGCCATCGTGTGTGCTCTCCGTCCTCTGCAATGGAGCTGGTGTGCTGAC 1806
QY 251 rValPheAlaGlyGlyProAlaProLeuProProValLysPheValValGlyAlaIleAl 271
Db 1807 CGTGTTCGTGGCGGGCTGCCCCCTGCCCCCGGTCAAGTTTGTGTAGTGGTGGATTGC 1866
QY 271 aGlyAlaAsnThrLeuThrGlyIleSerCysGlyLeuLeuAlaSerValAspAlaLeuTh 291
Db 1867 AGGCGCCAAACACCTTGACTGGCAATTCCTGTGGCTTCTAGCCTCAGTCTGATGCCCTGAC 1926
QY 291 rPheGlyGlnPheSerGluTyrGlyAlaArgTrpGluThrGlyLeuGlyCysArgAlaTh 311
Db 1927 CTTTGGTCAGTTCTCTGAGTACGGAGCCCGCTGGGAGACGGGGCTAGGCTCCCGGGCCAC 1986
```

Qy 311 rGlyPheLeuAlaValLeuGlySerGluAlaSerValLeuLeuLeuThrLeuAlaVa 331
 Db 1987 TGGCTTCTGGCAGTACTGGGTGGAGGACATCGGTGCTGCTGCTCACTCTGGCGCGAGT 2046
 Qy 331 lGlnCysSerValSerValSerCysValArgAlaTyrGlyLySerProSerLeuGlySe 351
 Db 2047 GCAGTGCAGCGTCTCCGCTCTCTGCTCGGGGCTATGGGAAGTCCCTCTCCCTGGGCG 2106
 Qy 351 rValArgAlaGlyValLeuGlyCysLeuAlaLeuAlaGlyLeuAlaAlaLeuProLe 371
 Db 2107 CGTTTCAGCAGGGGTCTTAGGTGCTGGCACTGGCAGGGTGGCGCGCACTGCCCT 2166
 Qy 371 uAlaSerValGlyGluTyrGlyAlaSerProLeuCysLeuProTyrAlaProProGluG1 391
 Db 2167 GGCCTCAGTGGAGATACGGGGCTCCCACTCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2286
 Qy 391 yGlnProAlaAlaLeuGlyPheThrValAlaLeuValMetMetAsnSerPheCysPheLe 411
 Db 2287 TCAGCCAGCAGCCCTGGGCTTCCACGCTGGCTTGGTATGATGAATCTCTTCTGTTTCT 2286
 Qy 411 uValValAlaGlyAlaTyrIleLeuLeuTyrCysAspLeuProArgGlyAspPheGluAl 431
 Db 2287 GGTGTCGTCGGCTGCTTACATCAAACTGTACTGTGACCTGGCGGGGCGACTTTGAGGC 2346
 Qy 431 aValTrpAspCysAlaMetValArgHisValAlaAlaTrpLeuIlePheAlaAspGlyLeuLe 451
 Db 2347 CGTGTGGGACTGGCGCATGGTGGAGCACGTGGCTGGCTCATCTTCGAGACGGGCTCT 2406
 Qy 451 uTyrCysProValAlaPheLeuSerPheAlaSerMetLeuGlyLeuPheProValThrPr 471
 Db 2407 CTACTGTCCCGTGGCTTCTCAGCTTCCCTCCATGCTGGGCTCTTCCCTGTGTCAGGCC 2466
 Qy 471 oGluAlaVallySerValLeuLeuValValLeuProLeuProAlaCysLeuAsnProLe 491
 Db 2467 CGAGGCCCTCAAGTCTGTCTGCTGGTGGTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGG 2526
 Qy 491 uLeuTyrLeuLeuPheAsnProHisPheArgAspAspLeuArgLeuArgProArgAl 511
 Db 2527 GCTGTACCTGCTCTTCAACCCCTTCCCGGATGACCTTCGGGGGCTTCGGGGGCTGG 2586
 Qy 511 aGlyAspSerGlyProLeuAlaTyrAlaAlaGlyLeuGlySerSerCysAs 531
 Db 2587 AGGGGACTCAGGGGCTTCTAGCTATGCTGGCGGGGAGCTGGAGAAGAGCTCTCTGTA 2646
 Qy 531 pSerThrGlnAlaLeuValAlaPheSerAspValAspLeuLeuLeuGluAlaSerGluAl 551
 Db 2647 TTCTACCCAGGCCCTGGTAGCTTCTCTGATGTGGATCTCATTTCTGGAAGCTTCTGAAGC 2706
 Qy 551 aGlyArgProProGlyLeuGluThrTyrGlyPheProSerValThrLeuIleSerCysG1 571
 Db 2707 TGGGGGGGCTTGGGCTGGAGACTATGCTTCCCTCAGTGACCTCTATCTCTCTGTA 2766
 Qy 571 nGlnProGlyAlaProArgLeuGlySerHisCysValGluProGluGlyAsnHisPh 591
 Db 2767 GCAGCCAGGGGCCCCAGGCTGGAGGCGAGCATTTGTGTAGAGCAGAGGGGACCACTT 2826
 Qy 591 eGlyAsnProGlnProSerMetAspGlyGluLeuLeuLeuArgAlaGluGlySerThrPr 611
 Db 2827 TGGGAACCCCAACCTTCCATGGATGGAGAACTGCTGTGAGGGCAGAGGGATCTAGGCC 2886
 Qy 611 oAlaGlyGlyLeuSerGlyGlyGlyPheGlnProSerGlyLeuAlaPheAlaSe 631
 Db 2887 AGCAGGTGGAGGCTTGTGAGGGGGTGGCGGCTTTTCAGCCCTCTGGCTTGGCTTGTCTC 2946
 Qy 631 rHisVal 633
 Db 2947 ACACGTG 2953

RESULT 15

US-10-331-496A-94
 ; Sequence 94, Application US/10331496A
 ; Publication No. US2003028305A1

GENERAL INFORMATION:

APPLICANT: PRANTZ, GRETCHE
 APPLICANT: HILLAN, KENNETH J.
 APPLICANT: PHILLIPS, HEIDI S.
 APPLICANT: POLAKIS, PAUL
 APPLICANT: SMITH, VICTORIA
 APPLICANT: SPENCER, SUSAN D.
 APPLICANT: WILLIAMS, P. MICKEY
 APPLICANT: WU, THOMAS D.
 APPLICANT: ZHANG, ZEMIN

TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE DIAGNOSIS AND
 TREATMENT OF TUMOR

FILE REFERENCE: P5014R1-PCT

CURRENT APPLICATION NUMBER: US/10/331,496A

CURRENT FILING DATE: 2002-12-30

PRIOR APPLICATION NUMBER: US 60/345,444

PRIOR FILING DATE: 2002-01-02

PRIOR APPLICATION NUMBER: US 60/351,885

PRIOR FILING DATE: 2002-01-25

PRIOR APPLICATION NUMBER: US 60/360,066

PRIOR FILING DATE: 2002-02-25

PRIOR APPLICATION NUMBER: US 60/362,004

PRIOR FILING DATE: 2002-03-05

PRIOR APPLICATION NUMBER: US 60/366,869

PRIOR FILING DATE: 2002-03-20

PRIOR APPLICATION NUMBER: US 60/366,284

PRIOR FILING DATE: 2002-03-21

PRIOR APPLICATION NUMBER: US 60/368,679

PRIOR FILING DATE: 2002-03-28

PRIOR APPLICATION NUMBER: US 60/404,809

PRIOR FILING DATE: 2002-08-19

PRIOR APPLICATION NUMBER: US 60/405,645

PRIOR FILING DATE: 2002-08-21

NUMBER OF SEQ ID NOS: 95

SEQ ID NO 94

LENGTH: 3443

TYPE: DNA

ORGANISM: Homo sapien

US-10-331-496A-94

Alignment Scores:

Pred. No.: 5,79e-308 Length: 3443

Score: 3045.50 Matches: 602

Percent Similarity: 92.61% Conservative: 12

Best Local Similarity: 90.80% Mismatches: 19

Query Match: 92.09% Indels: 32

DB: 17 Gaps: 4

US-10-664-667-5 (1-633) x US-10-331-496A-94 (1-3443)

Qy 1 AsnThrThrHisTyr----- 5

Db 971 AACTCCACACACTATCTCTGAATGGTGCCATGGACATCCAGGAGTTTCCAGATCTCAAAG 1030

Qy 6 ---ArgGluSerTrpTyrAla-----CysArgTyrArg 15

Db 1031 GCACCACCACCTGGAGATCTTACCTGACCCGCGCAGGATCCGGCTGCTCCCATCGG 1090

Qy 16 -----SerGlyIleProGlySer-ThrHisAlaSerValGluArgSerGln---G1 31

Db 1091 GGATGTGCCAACAGCTGCCAGGCTCGAGTCTCGGAACCTGTCTCACAATCAAATTCGAG 1150

Qy 31 YLeuSerLeuProAlaHisProAlaSerLeuAlaLeuAlaAlaSerAsnThrAl 51

Db 1151 AGCTGCCACCTGCACAGGTGTCAAAATTTGGAGGAAA-TCGGCCTTCCCAACAACCCG 1209

Qy 51 aSerGlyLyLeuGlu***AspThrPheSerGlnLeuSerSerLeuGlnAlaLeuAspLe 71

Db 1210 ATCTGGAAATTTGAGCT-GACACCTTCAGCCAGTGGCTCTCTGACCCCTTGGATCT 1268

Qy 71 uSerTrpAsnAlaIleArgSerIleHisProGluAlaPheSerThrLeuHisSerLeuVa 91

Db 1269 TAGCTGGAACGCCATCCGGTCCATCCACCTTGAGGCCTTCTCCACCCCTGCACCTCCCTGGT 1328

Qy 91 llysLeuAspLeuThrAspAenGlnLeuThrThrLeuProLeuAlaGlyLeuGlyGlyLe 111
Db 1329 CAAGCTGACCTGACAGACCAACAGCTGACACACTGCCCCCTGGCTGGACTTGGGGGCTT 1388
Qy 111 uMetHisLeuLysLeuLysGlyAsnLeuAlaLeuSerGlnAlaPheSerLysAspSerPh 131
Db 1389 GATGCATCTGAAGCTCAAAAGGGAACCTTGCTCTCCAGGCGCTTCTCAAGGACAGTTT 1448
Qy 131 eProLysLeuArgIleLeuGluValProTyrAlaTyrGlnCysCysProTyrGlyMetCy 151
Db 1449 CCAAAACTGAGATCTTGAGGTGCCCTTATGCCCTACAGTGCTGCTCCCTATGGATGTG 1508
Qy 151 sAlaSerPhePheLysAlaSerGlyGlnTrpGluAlaGluAspLeuHisLeuAspAspG1 171
Db 1509 TGCCAGCTTCTCAAGGCTCTCGGCGAGTGGAGGCTGAAGACCTTCACTTGTATGATGA 1568
Qy 171 uGlnSerSerLysArgProLeuGlyLeuLeuAlaArgGlnAlaGluAsnHisTyrAspG1 191
Db 1569 GGAGTCTTCAAAAAGGCGCCCTGGGCTCTCTTGGCCAGACAGAGAACCACTATGACCA 1628
Qy 191 nAspLeuAspGluLeuGlnLeuGluMetGluAspSerLysProHisProSerValGlnCy 211
Db 1629 GGACCTGATGAGCTCCAGCTGGAGATGGAGACTCAAAGCCACACCCAGTGTCCAGTG 1688
Qy 211 sSerProThrProGlyProPheLysProCysGluTyrLeuPheGluSerTrpGlyIleAr 231
Db 1689 TAGCCCTACTCCAGGCGCCCTTCAAGCCCTGTGAGTACCTCTTTGAAAGCTGGGGCATCG 1748
Qy 231 gLeuAlaValTrpAlaIleValLeuLeuSerValLeuCysAenGlyLeuValLeuLeuTh 251
Db 1749 CTTGGCCGTGGGCGCATCGTGTGCTCTCGTGTCTGCAATGAGCTGGTGTCTGCAC 1808
Qy 251 rValPheAlaGlyGlyProAlaProLeuProProValLysPheValValGlyAlaIleAl 271
Db 1809 CGTGTCTCGCGGCGCTGCCCGCTGCCCGGCTCAAGTTTGTGGTAGGTGCCATTGC 1868
Qy 271 aGlyAlaAsnThrLeuThrGlyIleSerCysGlyLeuLeuAlaSerValAspAlaLeuTh 291
Db 1869 AGGCGCAACACCTTGACTGGCATTTCTGTGGCCCTTCTAGCCCTCAGTCGATGCCCTGAC 1928
Qy 291 rPheGlyGlnPheSerGluTyrGlyAlaAArgTrpGluThrGlyLeuGlyCysArgAlaTh 311
Db 1929 CTTTGGTCAGTTCTCTGAGTAGCGAGCCGCTGGGAGACGGGGCTAGGCTGCCGGGCCAC 1988
Qy 311 rGlyPheLeuAlaValLeuGlySerGluAlaSerValLeuLeuLeuThrLeuAlaAlaVa 331
Db 1989 TGGCTTCTCGCAGTACTTGGGTGCGAGGCATCGGTGCTGCTGCTCACTCTGGCGGCAGT 2048
Qy 331 lGlnCysSerValSerValSerCysValAlaTyrGlyLysSerProSerLeuGlySe 351
Db 2049 GCAGTGCAGCGTCTCCGTCTCTGTGTCGGGCCCTATGGGAAGTCCCGCTCCCTGGGCAG 2108
Qy 351 rValArgAlaGlyValLeuGlyCysLeuAlaLeuAlaGlyLeuAlaAlaLeuProLe 371
Db 2109 CGTTTCAGCAGGGGTCTTAGGTGCTGGCACTGGCAGGGGTGGCGCGCACTGCCCT 2168
Qy 371 uAlaSerValGlyGluTyrGlyAlaSerProLeuCysLeuProTyrAlaProProGluG1 391
Db 2169 GGCTCTCAGTGGAGAAATACGGGGCTTCCCGCTCTGCTGCGCCCTACGCGCCACCTGAGGG 2228
Qy 391 yGlnProAlaAlaLeuGlyPheThrValAlaLeuValMetMetAenSerPheCysPheLe 411
Db 2229 TCAGCCAGACGCCCTGGGCTTCCCGTGGCCCTGGTGATGATGAATCTCTTCTGTTCCT 2288
Qy 411 uValValAlaGlyAlaTyrIleLysLeuTyrCysAspLeuProArgGlyAspPheGluAl 431
Db 2289 GGTCTGGCGCGTGCCTACATCAAACTGTACTGTGACCTGCGCGGGGCGACTTTGAGGC 2348
Qy 431 aValTrpAspCysAlaMetValArgHisValAlaTrpLeuIlePheAlaAspGlyLeuLe 451
Db 2349 CGTGTGGGACTGCGCGCATGGTGAGGCAGTGGCGCTGGCTCATCTTCGACAGACGGGCTCT 2408

Qy 451 uTyrCysProValAlaPheLeuSerPheAlaSerMetLeuLeuGlyLeuPheProValThrPr 471
Db 2409 CTACTGTCCCGTGGCCCTTCTCAGCTTCGCGCTCCATGCTGGGCGCTTCTCCCTGTCAACGCC 2468
Qy 471 oGluAlaValLysSerValLeuLeuValValLeuProLeuProAlaCysLeuAenProLe 491
Db 2469 CGAGGCGGTCAAGTCTGTCTGCTGGTGGTGTCTGCCCTGCTGCTGCTCAACCCACT 2528
Qy 491 uLeuTyrLeuLeuPheAenProHisPheArgAspAspLeuArgArgLeuArgProArgAl 511
Db 2529 GCTGTACTGTCTTCAACCCCGCTTCGCGGATGACCTTCGCGGCGCTTCGCGCGCGC 2588
Qy 511 aGlyAspSerGlyProLeuAlaTyrAlaAlaAlaGlyGluLeuGluLysSerSerCysAs 531
Db 2589 AGGGGACTCAGGGGCGCTTAGCTTATGCTGCGCGCGGGAGCTGGAGAAGAGCTCTCTGTGA 2648
Qy 531 pSerThrGlnAlaLeuValAlaPheSerAspValAspLeuIleLeuGluAlaSerGluAl 551
Db 2649 TTTACCCAGGCGCTTGTGAGCTTCTCTGATGTGGATCTCATCTTGGAAAGCTTCTGAAGC 2708
Qy 551 aGlyArgProProGlyLeuGluThrTyrGlyPheProSerValThrLeuIleSerCysG1 571
Db 2709 TGGGCGGCGCGCTGGGCTGGAGACTTATGGCTTCCCGCTCAGTGACCTCATCTCTGTCA 2768
Qy 571 nGlnProGlyAlaProArgLeuGluGlySerHisCysValGluProGluGlyAsnHisPh 591
Db 2769 GCAGCCAGGGGCGCGCTGAGGCTGGAGGCGAGCATTTGTGTAGAGCCAGAGGGAACCACTT 2828
Qy 591 eGlyAsnProGlnProSerMetAspGlyGluLeuLeuLeuArgAlaGluGlySerThrPr 611
Db 2829 TGGGAACCCCGCAACCCCTCCATGGATGGAGAACTGCTGCTGAGGGCAGAGGATCTACGCC 2888
Qy 611 oAlaGlyGlyGlyLeuSerGlyGlyGlyPheGlnProSerGlyLeuAlaPheAlaSe 631
Db 2889 AGCAGGTGGAGGCTTGTACGGGGTGGCGGCTTTCAGCCCTCTGGCTTGGCTTGTCTTC 2948
Qy 631 rHisVal 633
Db 2949 ACAGGTG 2955

Search completed: April 14, 2005, 22:03:29
Job time : 694 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 14, 2005, 11:59:23 ; Search time 219.132 Seconds
(without alignments)
14179.987 Million cell updates/sec

Title: US-10-664-667-6

Perfect score: 1899

Sequence: 1 aatacgaactactatagga.....ttggccttgcttcacacgtg 1899

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 1202784 seqs, 818138359 residues

Total number of hits satisfying chosen parameters: 2405568

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents_NA.*

- 1: /cgn2_6/ptodata/1/ina/5A_COMB.seq.*
- 2: /cgn2_6/ptodata/1/ina/5B_COMB.seq.*
- 3: /cgn2_6/ptodata/1/ina/6A_COMB.seq.*
- 4: /cgn2_6/ptodata/1/ina/6B_COMB.seq.*
- 5: /cgn2_6/ptodata/1/ina/ECTUS_COMB.seq.*
- 6: /cgn2_6/ptodata/1/ina/backfiles1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1752.8	92.3	2988	4	US-09-799-451-723
2	439.4	23.1	453	3	US-09-397-787-240
3	355.6	18.7	2724	4	US-09-170-496D-263
4	355.6	18.7	4570	4	US-09-976-594-201
5	352.4	18.6	2724	4	US-09-170-496D-277
6	194	10.2	2612	4	US-09-495-050A-214
7	194	10.2	4203	2	US-08-866-757-1
8	194	10.2	4203	3	US-09-153-593-1
9	181.8	9.6	723	4	US-09-495-050A-220
10	105	5.5	2703	4	US-09-482-273-75
11	101.8	5.4	473	2	US-08-866-757-3
12	101.8	5.4	473	3	US-09-153-593-3
13	94.6	5.0	2709	4	US-09-482-273-101
14	79.8	4.2	1898	4	US-09-589-510-7
15	78.8	4.1	570	2	US-08-633-682-1
16	78.8	4.1	570	3	US-08-936-772-1
17	78.8	4.1	570	3	US-09-395-918-1
18	78.4	4.1	1330	3	US-09-118-442-29
19	78.4	4.1	1330	3	US-09-677-064-29
20	77.8	4.1	1581	1	US-08-383-756-1
21	77.8	4.1	1581	2	US-08-460-898-1
22	77.6	4.1	1665	4	US-09-591-095-23
23	77.4	4.1	1390	4	US-09-205-258-124
24	77.4	4.1	4847	4	US-10-164-595-57
25	77	4.1	223	4	US-09-016-434-597
26	76	4.0	1297	4	US-09-800-729-80
27	76	4.0	1790	4	US-09-118-637A-3

28	75	3.9	3244	3	US-09-165-543-4	Sequence 4, Appli
29	73.8	3.9	1236	4	US-09-620-312D-1077	Sequence 1077, Ap
30	73.4	3.9	1236	4	US-09-620-312D-1077	Sequence 1077, Ap
31	69.2	3.6	1616	4	US-09-620-312D-1041	Sequence 1041, Ap
32	69.2	3.6	2019	3	US-09-063-950-3	Sequence 3, Appli
33	69.2	3.6	2852	3	US-09-063-950-1	Sequence 1, Appli
34	68	3.6	2765	4	US-09-620-312D-61	Sequence 61, Appli
35	66.8	3.5	2413	4	US-09-016-434-1469	Sequence 1469, Ap
36	66.8	3.5	3549	4	US-09-826-509-586	Sequence 586, App
37	65.8	3.5	1292	4	US-09-904-615-61	Sequence 1, Appli
38	65.6	3.5	538	1	US-08-890-572-1	Sequence 1, Appli
39	65.2	3.4	1597	3	US-09-038-832-3	Sequence 3, Appli
40	65.2	3.4	2292	4	US-09-826-509-394	Sequence 394, App
41	65.2	3.4	2292	4	US-09-826-509-418	Sequence 418, App
42	65.2	3.4	2292	4	US-09-826-509-422	Sequence 422, App
43	65	3.4	969	2	US-08-700-637-1	Sequence 1, Appli
44	64.8	3.4	831	4	US-09-904-615-25	Sequence 25, Appli
45	63.8	3.4	1262	4	US-09-489-847-69	Sequence 69, Appli

ALIGNMENTS

RESULT 1
US-09-799-451-723
; Sequence 723, Application US/09799451
; Patent No. 6783969
; GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom
; APPLICANT: Zhou, Ping
; APPLICANT: Goodrich, Ryle
; APPLICANT: Asundi, Vinod
; APPLICANT: Ren, Feiyan
; APPLICANT: Zhang, Jie
; APPLICANT: Xue, Aidong J.
; APPLICANT: Zhao, Qing A.
; APPLICANT: Wang, Jian-Rui
; APPLICANT: Ma, Yunqing
; APPLICANT: Yamazaki, Victoria
; APPLICANT: Chen, Rui-hong
; APPLICANT: Wang, Zhiwei
; APPLICANT: Wang, Dunrui
; APPLICANT: Yang, Yonghong
; APPLICANT: Wehrman, Tom
; APPLICANT: Ghosh, Reena
; APPLICANT: Drmanac, Radoje T.
; TITLE OF INVENTION: No. 6783969el Nucleic Acids and
; FILE REFERENCE: Polypeptides
; FILE REFERENCE: 803
; CURRENT APPLICATION NUMBER: US/09/799,451
; CURRENT FILING DATE: 2001-03-05
; NUMBER OF SEQ ID NOS: 948
; SOFTWARE: pt_fl_genes Version 2.0
; SEQ ID NO 723
; LENGTH: 2988
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(2484)
; US-09-799-451-723

Query Match					92.3%;	Score 1752.8;	DB 4;	Length 2988;
Best Local Similarity					99.8%;	Pred. No. 0;		
Matches 1765;					Conservative	0;	Mismatches	2;
							Indels	1;
							Gaps	1;
QY	132	AGCGGCTCCACACACCGCATCTGGAAATTTGGAGTNGACACCTTTCAGCCAGCTGAG	191					
Db	718	ATCGGCTCCACACACCGCATCTGGAAATTTGGAGCT-GACACCTTCAGCCAGCTGAG	776					
QY	192	CTCCCTGCAAGCCCTTGGATCTTAGCTGGAACCCATCCGGTCCATCCACCTTGGAGCCTT	251					
Db	777	CTCCCTGCAAGCCCTTGGATCTTAGCTGGAACCCATCCGGTCCATCCACCTTGGAGCCTT	836					

Qy 252 CTCACCCCTGCACCTCCCTGTCAGCTGAGCCTGACAGCAACACAGCTGACCACTGCC 311
 Db 837 CTCACCCCTGCACCTCCCTGTCAGCTGAGCCTGACAGCAACACAGCTGACCACTGCC 896
 Qy 312 CCTGGCTGGACTTGGGGCTTGGATGATCTGAAGCTCAAAGGGAACCTTGTCTCTCCCA 371
 Db 897 CCTGGCTGGACTTGGGGCTTGGATGATCTGAAGCTCAAAGGGAACCTTGTCTCTCCCA 956
 Qy 372 GGCCTTCTCCAGGACAGTTTCCCAAACTGAGGATCTGAGGATCTGAGGATCTTATGCTACCA 431
 Db 957 GGCCTTCTCCAGGACAGTTTCCCAAACTGAGGATCTGAGGATCTGAGGATCTTATGCTACCA 1016
 Qy 432 GTGCTGCTCCCTATGGGATGTGTCAGCTTCTCAAGGCTCTGGCAGTGGGAGGCTGA 491
 Db 1017 GTGCTGCTCCCTATGGGATGTGTCAGCTTCTCAAGGCTCTGGCAGTGGGAGGCTGA 1076
 Qy 492 AGACCTTCACTTGTATGATGAGGAGTCTTCAAAAGAGGCCCTTGGGCTCTTGGCCAGACA 551
 Db 1077 AGACCTTCACTTGTATGATGAGGAGTCTTCAAAAGAGGCCCTTGGGCTCTTGGCCAGACA 1136
 Qy 552 AGCAGAGAACCACTATGACCAAGGACCTGGATGAGCTCCAGCTGGAGATGGAGGACTCAAA 611
 Db 1137 AGCAGAGAACCACTATGACCAAGGACCTGGATGAGCTCCAGCTGGAGATGGAGGACTCAAA 1196
 Qy 612 GCCACACCCAGTGTCCAGTGTAGCCCTACTCCAGGCCCCCTTCAAGCCCTGTGAGTACT 671
 Db 1197 GCCACACCCAGTGTCCAGTGTAGCCCTACTCCAGGCCCCCTTCAAGCCCTGTGAGTACT 1256
 Qy 672 CTTTGAAGCTGGGGCATCCGCTGGCCGTGTGGCCATCGTGTGCTCTCCGTGCTCTG 731
 Db 1257 CTTTGAAGCTGGGGCATCCGCTGGCCGTGTGGCCATCGTGTGCTCTCCGTGCTCTG 1316
 Qy 732 CAATGGACTGTGTGCTGACCGTGTTCGCTGGGGGCTCCGCCCTTGGCCCCCGGTCAA 791
 Db 1317 CAATGGACTGTGTGCTGACCGTGTTCGCTGGGGGCTCCGCCCTTGGCCCCCGGTCAA 1376
 Qy 792 GTTTGTGATGAGTGGGATGACGGGCCCAACACTTGAAGTGGGATTTCTGTGGCTTCT 851
 Db 1377 GTTTGTGATGAGTGGGATGACGGGCCCAACACTTGAAGTGGGATTTCTGTGGCTTCT 1436
 Qy 852 AGCCTCAGTCAGTCCCTGACCTTTGGTCAGTCTCTGAGTACGAGGCCCTGGAGAC 911
 Db 1437 AGCCTCAGTCAGTCCCTGACCTTTGGTCAGTCTCTGAGTACGAGGCCCTGGAGAC 1496
 Qy 912 GGGGTAGGCTGCGGGCCACTGGCTTCTGGCAGTACTTGGGTGGAGGCAATCGGTGCT 971
 Db 1497 GGGGTAGGCTGCGGGCCACTGGCTTCTGGCAGTACTTGGGTGGAGGCAATCGGTGCT 1556
 Qy 972 GCTGCTACTTGGCCGAGTGCAGTGCAGGCTCTCCGTCTCCTGTGTCGGGCTATGG 1031
 Db 1557 GCTGCTACTTGGCCGAGTGCAGTGCAGGCTCTCCGTCTCCTGTGTCGGGCTATGG 1616
 Qy 1032 GAAGTCCCTCCCTGGGACGCTTCCAGAGGAGGCTCTAGGCTGCTGGGCACTGGCAGG 1091
 Db 1617 GAAGTCCCTCCCTGGGACGCTTCCAGAGGAGGCTCTAGGCTGCTGGGCACTGGCAGG 1676
 Qy 1092 GCTGGCCGCGCACTGCCCTTGGGCTCAGTGGGAGAAATACGGGGCTTCCCACTCTGCT 1151
 Db 1677 GCTGGCCGCGCACTGCCCTTGGGCTCAGTGGGAGAAATACGGGGCTTCCCACTCTGCT 1736
 Qy 1152 GCCCTACGCGCCACTGAGGCTCAGCCAGCCCTGGGCTTCAACGTTGGCCCTGGTGAT 1211
 Db 1737 GCCCTACGCGCCACTGAGGCTCAGCCAGCCCTGGGCTTCAACGTTGGCCCTGGTGAT 1796
 Qy 1212 GATGAATCTCTTCTGTTTCTGTTGCTGGGCGGCTTACATCAAACTGACTGTGACCT 1271
 Db 1797 GATGAATCTCTTCTGTTTCTGTTGCTGGGCGGCTTACATCAAACTGACTGTGACCT 1856
 Qy 1272 GCCCGGGGCGACTTTGAGGCGGCTGTGGGACTGCGCCATGCTGAGGCACTGGGCTTGCT 1331
 Db 1857 GCCCGGGGCGACTTTGAGGCGGCTGTGGGACTGCGCCATGCTGAGGCACTGGGCTTGCT 1916

Qy 1332 CATCTTCGACAGGGGCTCTTACTGTCCGTGGCCCTTCTCCTCAGCTTCGCTCCATGCT 1391
 Db 1917 CATCTTCGACAGGGGCTCTTACTGTCCGTGGCCCTTCTCCTCAGCTTCGCTCCATGCT 1976
 Qy 1392 GGGCTCTTCTCCTGTACGCCCGGAGGCGCTCAAGTCTGTCTCCTGTGTGTGCTGCCCT 1451
 Db 1977 GGGCTCTTCTCCTGTACGCCCGGAGGCGCTCAAGTCTGTCTCCTGTGTGTGCTGCCCT 2036
 Qy 1452 GCTGCTGCTCTCAACCCACTGTGTACTGTCTTCAACCCCACTTCCGGGATGACCT 1511
 Db 2037 GCTGCTGCTCTCAACCCACTGTGTACTGTCTTCAACCCCACTTCCGGGATGACCT 2096
 Qy 1512 TCGGCGGCTTCGGCCCGCGCAGGGGACTCAGGSCCTTAGCCTATGCTGGGCGGGGA 1571
 Db 2097 TCGGCGGCTTCGGCCCGCGCAGGGGACTCAGGSCCTTAGCCTATGCTGGGCGGGGA 2156
 Qy 1572 GCTGGAGAAGAGCTCTCTGTGATTTACCCAGGCCCTGTGAGCTTCTCTGATGTGGATCT 1631
 Db 2157 GCTGGAGAAGAGCTCTCTGTGATTTACCCAGGCCCTGTGAGCTTCTCTGATGTGGATCT 2216
 Qy 1632 CATCTGGAAGCTTCTGAAGCTGGGCGGCCCTTGGGCTGGAGACCTATGCTTCCCTC 1691
 Db 2217 CATCTGGAAGCTTCTGAAGCTGGGCGGCCCTTGGGCTGGAGACCTATGCTTCCCTC 2276
 Qy 1692 AGTGACCTCATCTCTCTGTGAGCAGCAGGCGGCCCTCAGGCTGGAGGAGGAGGCTTGT 1751
 Db 2277 AGTGACCTCATCTCTCTGTGAGCAGCAGGCGGCCCTCAGGCTGGAGGAGGAGGCTTGT 2336
 Qy 1752 AGAGCCAGAGGGAACCACTTTGGGAACCCCAACCTTCCATGATGGAGAACTGCTGCT 1811
 Db 2337 AGAGCCAGAGGGAACCACTTTGGGAACCCCAACCTTCCATGATGGAGAACTGCTGCT 2396
 Qy 1812 GAGGCGAGAGGATCTACGCCAGCAGGTGAGGCTTGTACGGGGGTGGGCTTTCAGCC 1871
 Db 2397 GAGGCGAGAGGATCTACGCCAGCAGGTGAGGCTTGTACGGGGGTGGGCTTTCAGCC 2456
 Qy 1872 CTCTGGCTTCGCTTTCACACGTTG 1899
 Db 2457 CTCTGGCTTCGCTTTCACACGTTG 2484

RESULT 2
 US-09-397-787-240
 ; Sequence 240, Application US/09397787
 ; Patent No. 6468758
 ; GENERAL INFORMATION:
 ; APPLICANT: Benson, Darin R.
 ; APPLICANT: Lodes, Michael J.
 ; APPLICANT: Mitcham, Jennifer L.
 ; APPLICANT: King, Gordon E.
 ; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR OVARIAN
 ; FILE REFERENCE: 210121.466C2
 ; CURRENT APPLICATION NUMBER: US/09/397,787
 ; CURRENT FILING DATE: 1999-09-16
 ; NUMBER OF SEQ ID NOS: 334
 ; SOFTWARE: FastSeq for Windows Version 3.0
 ; SEQ ID NO 240
 ; LENGTH: 453
 ; TYPE: DNA
 ; ORGANISM: Homo sapien
 ; US-09-397-787-240

Query Match 23.1%; Score 439.4; DB 3; Length 453;
 Best Local Similarity 99.6%; Pred. No. 5.8e-92;
 Matches 451; Conservative 0; Mismatches 1; Indels 1; Gaps 1;
 Qy 1328 GGCTCATCTTCGACAGCGGCTCTTACTGTCCGTGGCTTCTCCTCAGCTTCGCTTCCA 1387
 Db 1 GGCTCATCTTCGACAGCGGCTCTTACTGTCCGTGGCTTCTCCTCAGCTTCGCTTCCA 60
 Qy 1388 TGCTGGGCTCTTCCCTGTGACGCCGAGGCGCTCAAGTCTGCTCCTGCTGCTGCTGC 1447

```
Db 61 TGCTGGGCGCTCTTCCCTGTCTACGCCCGGCGGTCAAGTCTGTCTGTCTGTGTGTGTGTGC 120
Qy 1448 CCCTCCCTGCTCCCTCAACCCACCTGCTGTACCTCTCTTCAACCCCCACCTTCCGGGATG 1507
Db 121 CCCTGCTGCTGCTCAACCCACCTGCTGTACCTCTCTTCAACCCCCACCTTCCGGGATG 180
Qy 1508 ACCTTCGGCGGCTTCGGCGCCCGCGCAGGGGACTCAGGGCCCTTACGGCCCTTACGCTATGCTGCGCGG 1567
Db 181 ACCTTCGGCGGCTTCGGCGCCCGCGCAGGGGACTCAGGGCCCTTACGGCCCTTACGCTATGCTGCGCGG 240
Qy 1568 GGGAGCTGGAGAGAGCTCCTGTGATCTACCCAGGCCCTTGTGAGCTTCTCTGATGTGG 1627
Db 241 GGGAGCTGGAGAGAGCTCCTGTGATCTACCCAGGCCCTTGTGAGCTTCTCTGATGTGG 300
Qy 1628 ATCTCATTTCTGGAAGCTTCTGAAGCTGGCGGCCCTTGGGCTGGAGACCTATGGCTTCC 1687
Db 301 ATCTCATTTCTGGAAGCTTCTGAAGCTGGCGGCCCTTGGGCTGGAGACCTATGGCTTCC 360
Qy 1688 CCTCAGTGACCTCATCTCTGTCTAGCAGCCAGGGGCCCTCAGGGCTGGAGGCG-AGCCAT 1746
Db 361 CCTCAGTGACCTCATCTCTGTCTAGCAGCCAGGGGCCCTCAGGGCTGGAGGCGAAGCCAT 420
Qy 1747 TGTGTAGAGCCAGAGGGAACCACTTTTGGGAAC 1779
Db 421 TGTGTAGAGCCAGAGGGAACCACTTTTGGGAAC 453

RESULT 3
US-09-170-496D-263
; Sequence 263, Application US/09170496D
; Patent No. 6555339
; GENERAL INFORMATION:
; APPLICANT: Behan, Dominic P.
; APPLICANT: Chalmers, Derek T.
; APPLICANT: Liaw, Chen W.
; TITLE OF INVENTION: No. 6555339-Endogenous, Constitutively Activated Human G Protein-
; TITLE OF INVENTION: Receptors
; FILE REFERENCE: AREN-0040
; CURRENT APPLICATION NUMBER: US/09/170,496D
; CURRENT FILING DATE: 1998-10-13
; NUMBER OF SEQ ID NOS: 294
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 263
; LENGTH: 2724
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-170-496D-263

Query Match 18.7%; Score 355.6; DB 4; Length 2724;
Best Local Similarity 56.2%; Pred. No. 2.1e-72;
Matches 781; Conservative 0; Mismatches 580; Indels 28; Gaps 5;

Qy 135 GGCCTCCAAACACACCGCATCTGGAAATTTGGAGCTNGACACCTTACGCCAGCTGAGCTC 194
Db 1138 GACCTTAAGACATAATGAAATCTACGAAATTTAAAG-TTGACACATTTCCAGCAGTTGCTTAG 1196
Qy 195 CCTCAAGCCCTGGATCTTAGCTGAACGCCATCCGGTCCATCCACCTGAGGCCCTCTC 254
Db 1197 CCTCGATCGCTGAATTTGGCTTGGAAACAAATTTGCTATTATTCACCCCAATGCAATTTTC 1256
Qy 255 CACCTCGACCTCCCTGGTCAAGCTGGAGCTTGACAGACCAACAGCTGACCACTGCCCT 314
Db 1257 CACTTTGGCATCCCTAATAAGCTGAGCTATCGTCCAACTCCCTGCTGCTTTTCCCTAT 1316
Qy 315 GGCTGGACTTGGGGCTTGATGATCTGAAGCTCAAAGGAAACCTTGTCTCTCCAGGC 374
Db 1317 AACTGGGTTACATGTTTAACTCACTTAAATTTAAGGAAATTAACAGGAAATCATGCTTACAGAGCTT 1376
Qy 375 CTTCTCCAAGACAGATTTCCAAACTGAGGATCTCTGGAGGTGCTTATGCTACCACTG 434
Db 1377 GATATCATCTGAAACATTTTCCAGAACTCAAGGTTATAGAAATGCTTATGCTTACCACTG 1436
Qy 435 CTGTCCCTATGGGATGTGTGCCAGCTTCTTCAAGGCCCTCTGGGAGCTGGAGGCTGAAGA 494
```

```
Db 1437 CTGTGCATTTTGGAGTGTGTGAGAAATGCCTATAAGATTTCTAATCAATGGAATAAAGGTGA 1496
Qy 495 CTTTACCTTGATGATGAGGAGTCTTCAAAAGGCCCTCGGCCCTCTTGCAGACAAAGC 554
Db 1497 CAACAGCAGTATGAGCAGACCTTCAAGAAA-----GATGCTGGATGTT 1541
Qy 555 AGAGAACCACTATACACAGGACCTGGATGAGCTCCAGCTGGAGAT---GGAGGACTCAAA 611
Db 1542 TCAGGCTCAAGATGAACGTGACCTTGAAGATTTCTGCTTGAATTTGAGGAAAGACCTGAA 1601
Qy 612 GCCACACCCAGTGTCAAGTGTAGCCCTACTCCAGGCCCTTCAAGCCCTGTGAGTACT 671
Db 1602 AGCCCTTCAATTCAGTGTGAGTGTACCTTCCAGGCCCTTCAAGCCCTGTGAAACCT 1661
Qy 672 CTTTGAAGCTTGGGGCATCCGCTGGCGTGTGGGCATCGTGTGTGCTCTCCGTGCTCTG 731
Db 1662 GCTTGTATGGCTGGCTGTGATCAGAAATGGAGTGTGACCATAGCAGTTCGTGGCATTACTTG 1721
Qy 732 CAATGGACTGCTGTGCTGACCGTGTGTGGTGGCGGCTGCCGCCCTTGCCTCCCGGTCAA 791
Db 1722 TAATGCTTTGGTGACTTCAACAGTTTTTCAAGATCC---CCTCTGTACATTTCCCTCATTA 1778
Qy 792 GTTTGTGTAGTGGGATTCAGAGCGGCCCAACACTTGAATGGCATTTCTGTGGCTTCT 851
Db 1779 ACTGTTAATTTGGGGTTCATCGCAGCAGTGAACATGCTCAGCGGAGTCTCCAGTGGCGTCT 1838
Qy 852 AGCCTCAGTCAGTGCCTGACCTTTTGTGTGCTCTCTGAGTACGAGGCCCTGGAGAC 911
Db 1839 GGCTGGTGTGATGCGTTTCACTTTTGGCAGCTTTGACGACATGCTGCTGGTGGAGAA 1898
Qy 912 GGGCTAGGCTGCCGGGCCACTTGGCTTCTCTGGCAGTACTTGGTGGGAGGATCGGTGCT 971
Db 1899 TGGGTTGGTTGCCATGTGCTTGTGTTTGTGCTTTCATTTTCTTTCAGAAATCATCTGTTT 1958
Qy 972 GCTGCTCATCTGCGCGCAGTGTGAGTGCAGCGTCTCGTCTCTGCTGTGCTGGGCCCTATGG 1031
Db 1959 CCTGCTTACTCTGGCAGCCCTGGAGCGTGGTGTCTCTGTGAAATATTCTGCAAAATTTGA 2018
Qy 1032 GAAGTCCCTCCCTCGTGGCAGCGTTCGAGCAGGGGTCTTAGGCTGCCCTGGCACATGGCAGG 1091
Db 2019 AACGAAAGCTCCATTTTCTAGCCTGAAAGTAAATCATTTTGTCTGTGCCCTGTGGCCTT 2078
Qy 1092 GCTGGCGCCGCACTGCCCTCTGGCTCTCAGTGGGAGAAATACGGGGCTCTCCCACTCTGCT 1151
Db 2079 GACCATGGCGCAGTTCCTCTGCTGGGTGGCAGCAAGTATGCGGCTCTCCCTCTCTGCT 2138
Qy 1152 GCCCTAGCGGCCACTGAGGCTCAGCCAGCAGCCCTGGGCTTTCACCGTGGGCCCTGGGTAT 1211
Db 2139 GCCTT-----TGCTTTTGGGAGCCCGCAGCACCATGGGCTTACATGGTCTCATCTT 2192
Qy 1212 GATGAATCTCTTCTGTTTCCGTGCGTGGCGGCTTACATCAAACTGTACTGTGACCT 1271
Db 2193 GCTCAATTTCCCTTTGCTTCTCATGATGACCATTTGCTTACACCAAGCTCTACTGCAATTT 2252
Qy 1272 GCCCGGGGGCACTTTTGGGCGCTGTGGGACTGGGCCATGTGTGAGGACGCTGGCCCTGCT 1331
Db 2253 GGACAAAGGAGACCTGGAGAAATTTGGGACTGCTATGTTAAACACATTTGCCCTGTT 2312
Qy 1332 CATTTTGGCAGAGGGGCTCTCTACTGTCTCGTGGCGCTTCTCTGAGCTTGGCTTCCATGCT 1391
Db 2313 GCTCTTACCACAACTGCATCTTAAACTGCCCTGTGGCTTCTTGTCTCTCTCTCTTAAAT 2372
Qy 1392 GGGCTCTTCTCTGTCAGGCCCGGCGGTCAAGTCTGTCTGTGCTGGTGGTGGTGGCTCCCT 1451
Db 2373 AAACCTTTACATTTATCAGTCTGAAATTAAGTTTATCTCTTCTGTGGTGGTAGTCCCACT 2432
Qy 1452 GCCTGCTGCTCAACCCACTGTGTACTGTCTTCAACCCCACTTCCGGGATGACCT 1511
Db 2433 TCCTGATGCTCAATCCCTTCTCTACATCTTGTTCATCTCTCACTTTAAGGAGGATCT 2492
Qy 1512 TCGGCGGCT 1520
```

Db 2493 GGTGAGCCT 2501

RESULT 4

US-09-976-594-201
; Sequence 201, Application US/09976594
; Patent No. 6673549
; GENERAL INFORMATION:

; APPLICANT: Furness, Michael
; APPLICANT: Buchbinder, Jenny
; TITLE OF INVENTION: GENES EXPRESSED IN C3A LIVER CELL CULTURES TREATED WITH STEROIDS
; FILE REFERENCE: PA-0041 US
; CURRENT APPLICATION NUMBER: US/09/976,594
; PRIOR FILING DATE: 2001-10-12
; PRIOR APPLICATION NUMBER: 60/240,409
; PRIOR FILING DATE: 2000-10-12
; NUMBER OF SEQ ID NOS: 1143
; SOFTWARE: PERL Program
; SEQ ID NO 201
; LENGTH: 4570
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No. 6673549 205542.2
; NAME/KEY: unsure
; LOCATION: 3900, 3919, 3934
; OTHER INFORMATION: a, t, c, g, or other
US-09-976-594-201

Query Match 18.7%; Score 355.6; DB 4; Length 4570;
Best Local Similarity 56.2%; Pred. No. 2,4e-72;
Matches 781; Conservative 0; Mismatches 580; Indels 28; Gaps 5;

Qy 135 GGCCTCCACACACCCGATCTGGGAAATGGAGCTNGACACCTTCAGCCAGCTGAGCTC 194
Db 1399 GACCTTAAGACATAATGAAATCTACGAAATTAAG-TTGACACTTTCAGCAGCTTGTAG 1457
Qy 195 CCTGCAAGCCCTGGATCTAGCTGGAAGCCATCGGTCCATCCAGCTGAGGCTTCTC 254
Db 1458 CCTCGATCGCTGAAATTTGGCTTGGAAACAAATTTGCTAATTCACCCCAATGATTTTC 1517
Qy 255 CACCCTGCACCTCCCTGGTCAAGCTGGACCTGACAGACAAACAGCTGACACACTGCCCT 314
Db 1518 CACTTTGCCATCCCTAATAAGCTGGACCTATCGTCCAACTCTCTGCTTTCTCTAT 1577
Qy 315 GGCTGACCTTGGGGCTTGTGATGCACTGGAAGCTCAAGGAAACCTTGTCTCTCCAGGC 374
Db 1578 AACTGGGTTACATGGTTTAACTCACTTAAATTAACAGGAAATCATGCTTACAGAGCTT 1637
Qy 375 CTCTCCAAAGACAGCTTCCCAAACTGAGGATCTGGAGTGGCTTATGCTTACCACTG 434
Db 1638 GATATCATCTGAAACCTTCCAGAACTCAAGGTTATAGAAATGCTTATGCTTACCACTG 1697
Qy 435 CTGTCCCTATGGGATGTGTGCAGCTCTTCAAGCCCTCTGGCAGTGGAGGCTGAAGA 494
Db 1698 CTGTGCAATTTGGAGTGTGAGATGCTATAGATTCTAATCAATGGAATAAGGTGA 1757
Qy 495 CCTTCACTTATGATGAGGAGTCTTCAAAAGGCCCTGGGCCCTCTTGGCAGACAGC 554
Db 1758 CAACAGCAGTATGACGACCTTCATAAGAAA-----GATGCTGAATGTT 1802
Qy 555 AGAGAACCACTATGACAGGACCTGGATGAGCTCCAGCTGGAGAT---GGAGGACTCAA 611
Db 1803 TCAGGCTCAAGATGAACGATGACCTTGAAGATTTCTGCTTGAATTTGAGGAAAGACCTGAA 1862
Qy 612 GCCACACCCAGTGTCCAGTGTAGCCCTTACCTCCAGGCCCTTCAAGCCCTGTAGTACCT 671
Db 1863 AGCCCTTCATTCAGTGCAGTGTTCACCTTCCAGGCCCTTCAAACTGTGACACT 1922
Qy 672 CTTTGAAGCTGGGGCATTCGCCCTGGCGGTGGGCCCATGCTGTGCTCTCCGTGCTG 731
Db 1923 GCTTGTGCTGGCTGTGATCAGAAATGGAGTGTGGACCATAGCAGTTCTGGCACTTACTTG 1982

Qy 732 CAATGGAGCTGCTGCTGACCCGTTTCGCTGGGGGCTGCCCCCTGCCCCCGGTCAA 791
Db 1983 TAATGCTTTGGTGACTTCAACAGTTTTTCAATCC---CCTCTGTACATTTCCCAATTA 2039
Qy 792 GTTTGTGAGGTGCGATTCAGCGCCCAACACCTTGACTGGCATTTCTGTGACCTTCT 851
Db 2040 ACTGTTAATTTGGGTGCTATCGCAGCAGTGAACATGCTCAGCGAGTCTCCAGTCCGTGCT 2099
Qy 852 AGCCTCAGTGTATGCCCTGACCTTTTGGTCAAGTTCTCTGAGTACGGAGCCCTGGAGAC 911
Db 2100 GGCTGTGTGATGCTGCTTTCATTTTGGCAGCTTTTGCACGACATGCTGCTGGGAGAA 2159
Qy 912 GGGCTAGGCTGCGCGGCACCTGCTTCTCGGAGTACTTGGGTGCGAGGACATCGGTGCT 971
Db 2160 TGGGTTGGTTGGCATGTCATGTTTTTGTTCATTTTGTCTCTGCTCAGAAATCATCTGTTT 2219
Qy 972 GCTGCTCACTTGGCGCAGTGCAGCTGCTCCGCTCTCTCTCTCTCTCTCTCTCTCTCT 1031
Db 2220 CCTGCTTACTTGGCAGCCCTGGAGCGTGGGTTCTCTGTGAAATATTTCTGCAAAATTTGA 2279
Qy 1032 GAAGTCCCTTCCCTGGGACGCTTCGAGCAGGGGTCTTAGGCTGCTGACACTGGCAGG 1091
Db 2280 AACGAAAGCTCCATTTTCTAGCCCTGAAAGTAATCAATTTGCTCTGTGCGCTTGTGGCCT 2339
Qy 1092 GCTGGCGCGCAGCTGCTCCCTGCGCTCAGTGGGAGAAATACGGGCGCTCCCGACCTCTGCT 1151
Db 2340 GACCATGGCGCAGTTCCTCTGCTGGTGGCAGCAAGTATGGCGCCTCCCTCTCTCTGCT 2399
Qy 1152 GCCCTACGCGCCACCTGAGGGTTCAGCCAGCAGCCCTGGGCTTCACCGTGGGCCCTGTGAT 1211
Db 2400 GCCTT-----TGCTTTTGGGAGCCCGACACCATGGCTACATGCTGCTCATCTT 2453
Qy 1212 GATGAATCTCTTCTGTTTCTGTTGCTGGCGGTGCTTACATCAAACTGTACTGTGACCT 1271
Db 2454 GCTCAATTCCTTTGCTTCTCATGATGACCAATTTGCTTACACCAAGCTCTACTGCAATTT 2513
Qy 1272 GCCGCGGGGACCTTTGAGGCGCTGTTGGGAGTGGCCATGCTGAGGACGCTGGCGCTGCT 1331
Db 2514 GGACAGGAGACCTGGAGAAATTTGGGAGTGTCTATGTTAAACACATGTCCTGTT 2573
Qy 1332 CATCTTCGAGAGGGCTCTCTACTGTCTCCGTTGGCTTCTCTCAGCTTTCGCTCCATGCT 1391
Db 2574 GCTCTTCAACCACTGCTATCTAACTGCTGCTGCTTCTTGTCTCTCTCTCTTAAAT 2633
Qy 1392 GGGCTCTTCTCTGTCAGCGCCGAGGCGCTCAAGTCTGCTGCTGTGCTGCTGCTGCT 1451
Db 2634 AAACCTTACATTTATCAGTCTCGAAAGTAATTAAGTTTATCTCTGCTGCTGCTGCT 2693
Qy 1452 GCCTGCTGCTCAACCCACTGCTGCTGCTCTTCAACCCCGACCTTCCGGGATGACCT 1511
Db 2694 TCCTGCAATGCTCAATCCCTTCTCTACATCTTGTTCATCTCTCTCTCTCTCTCTCT 2753
Qy 1512 TCGGCGGCT 1520
Db 2754 GGTGAGCCT 2762

RESULT 5

US-09-170-496D-277
; Sequence 277, Application US/09170496D
; Patent No. 6555339
; GENERAL INFORMATION:
; APPLICANT: Behan, Dominic P.
; APPLICANT: Chalmers, Derek T.
; APPLICANT: Liaw, Chen W.
; TITLE OF INVENTION: No. 6555339-Endogenous, Constitutively Activated Human G Protein-
; FILE REFERENCE: AREN-0040
; CURRENT APPLICATION NUMBER: US/09/170,496D
; CURRENT FILING DATE: 1998-10-13
; NUMBER OF SEQ ID NOS: 294
; SOFTWARE: PatentIn version 3.1

```
; SEQ ID NO 277
; LENGTH: 2724
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-170-496D-277

Query Match      18.6%; Score 352.4; DB 4; Length 2724;
Best Local Similarity 56.1%; Pred. No. 1.1e-71;
Matches 779; Conservative 0; Mismatches 582; Indels 28; Gaps 5;

QY 135 GGCCTCCACACACCGCATCTGGGAATTTGGAGCTNGACACCTTCAGCCAGCTGAGTCTC 194
Db 1138 GACCTTAAGACATAATGAAATCTACGAAATTTAAAG-TTGACATTTTCCAGCAGTTGCTTAG 1196
QY 195 CCTCAAGCCCTGGATCTTAGCTGAAAGCCATCGGTCATCCACCCCTGAGGCCCTTCTC 254
Db 1197 CCTCGATCGTGAATTTGGCTTGGACAAATTTGCTATTATCACCCCAATGCATTTTC 1256
QY 255 CACCCTGACCTCCCTGGTCAAGCTGGAAGCTGACAGACCAACAGCTGACCAACATGCCCT 314
Db 1257 CACTTTGCCATCCCTTAATAAAGCTGGACCTATCGTCCAACCTCTGCTGCTTTTCTAT 1316
QY 315 GGCTGGACTTGGGGCTTGGATCTGACATCTGACCTCAAGGGRACCTTGTCTCTCCAGGC 374
Db 1317 AACTGGGTTACATGTTTAACTCACTTAAATTAACAGGAATCATGCTTACAGAGCTT 1376
QY 375 CTTCTCCAAGGACAGTTTCCCAAACTGAGGATCTCGAGGTGCTTATGCTTACCACTG 434
Db 1377 GATATCATCTGNAACCTTCCAGAACTCAAGTTATAGAAATGCTTATGCTTACCACTG 1436
QY 435 CTGCTCCCTATGGATGTGTGCAGCTTCTTCAAGCCCTCTGGGAGTGGAGGCTGAAGA 494
Db 1437 CTGTGCAATTTGGAGTGTGTGAGATGCTTAAAGATTTCTAATCAATGAATAAAGTGA 1496
QY 495 CTTTCACCTTGNATGATGAGGAGCTTCAAAAGGCCCTCGGCCCTCTTCCAGACAAGC 554
Db 1497 CAACAGCAGTATGGACGACCTTCAATGAANA-----GATGTGGAATGTT 1541
QY 555 AGAGAACCACTATGACAGGACCTGGATGAGCTCCAGCTGGAGAT---GGAGGACTCAA 611
Db 1542 TCAGGCTCAAGATGAACGTGACCTTGAAGATTTCTGCTGATTTGAGGAAGACCTGAA 1601
QY 612 GCCACACCCAGTGTCCAGTGTAGCCCTTACTCCAGGCCCTTCAAGCCCTGTGAGTACCT 671
Db 1602 AGCCCTTCATTCACTGAGTGTACCTTCCAGGCCCTTCAAAACCTGTGAACACCT 1661
QY 672 CTTTGAAGCTGGGGCATCCGCTGGCGTGTGGGCCATGCTGTGCTCTCGCTGCTG 731
Db 1662 GCTTGATGGCTGGCTGATCAGAAATTTGGAGTGTGGACCATAGCAGTTCTGGCATTACTG 1721
QY 732 CAATGGACTGTGTGCTGACCGTGTTCGCTGGGGGCTGCCCCCTGCCCGGTCAA 791
Db 1722 TAATGCTTTGGTGTACTTCAACAGTTTTTCAATCC---CCTGTGACATTTCCCAATTAA 1778
QY 792 GTTGTGTAGTGTGGATTTGAGCGGCCCAACACCTTGTGATGCTGATTTCTGTGGCTTCT 851
Db 1779 ACTGTTAATTTGGGTTCATCGCAGCAGTGAACATGCTCAGCGGAGTCTCCAGTGGCGTCT 1838
QY 852 AGCCTCAGTCGATGCCCTGACCTTTGGTGTGAGTCTCTGAGTACGAGGCCGCTGGGAGAC 911
Db 1839 GGCTGGTGTGATGCTGTTCACTTTTGGGAGCTTTGACAGCATGTTGCTGCTGGTGGAGAA 1898
QY 912 GGGGCTAGCTGCCGGGCCACTGGCTCTCTGACGACTTGGTTCGGAGGATCGGTGCT 971
Db 1899 TGGGGTTGGTGGCATGTCATTTGGTTTTTGTGTCATTTTGTCTTCAGAAATCATCTGTTTT 1958
QY 972 GCTGCTACTCTGGCGGAGTGTGAGCGTCTCTCGTCTCTGCTGTGCTCGGGCCCTATGG 1031
Db 1959 CTGCTTACTCTGGCAGCCCTGGACGCTGGGGTCTCTGTGAAATATTCTGCAGAAATTTGA 2018
QY 1032 GAATFCCCCCTCCCTGGGGAGCGTTCGAGCAGGGGTCTTAGCTGCCCTGGCAGCTGGCAGG 1091
Db 2019 AACGAAAGCTCCATTTTCTAGCCTGAAAGTAATCATTTTTTGTCTCTGTGCCCTGCTGGCCTT 2078
```

```
QY 1092 GCTGGCGCGCGCACTTGCCCTTGCGCTCAGTGGGAGAAATACGGGGCTCCCACTCTGCT 1151
Db 2079 GACCATGGCGCGAGTTCCCTTGCTGGGTGGGAGCAAGTATGGCGCTCCCTCTCTGCT 2138
QY 1152 GCCCTACGGCGCACTGAGGGTCAAGCAGAGCCCTGGGCTTCAAGTGGGCCCTGGTGA 1211
Db 2139 GCCTT-----TGCTTTTGGGAGCCAGCACCATGGGCTACATGGTCTCATCTT 2192
QY 1212 GATGAATCTCTTCTGTTTCTGGTCTGGCGGCTGCTTACATCAAACTGTACTGTGACCT 1271
Db 2193 GCTCAATTTCCCTTGTCTTCTCATGATGACATTTGCCCTACACCAAGCTCTACTGCAATTT 2252
QY 1272 GCCGGCGGCGCACTTTGAGGCGCTGTGGGACTGGGCCATGCTGAGGCAAGCTGGCCCTG 1331
Db 2253 GGACAAGGGAGACCTGGAGATATTTGGGACTGCTATGAAAAAACAACATTTGCCCTGTT 2312
QY 1332 CATCTTGCAGACGGGCTCTTACTGCTCCGCTGGGCTTCTCAGCTTGGCTTCCATGCT 1391
Db 2313 GCTCTTCAACCAACTGCATCTCTAAACTGCCCTGTGGCTTTCTTGTCTCTCTCTTTAAT 2372
QY 1392 GGGCTCTTCCCTGTCAGCGCCGAGCGCTCAAGTCTGCTGCTGGTGGTGGTGGCCCT 1451
Db 2373 AAACCTTACATTTATCAGTCTCTGAAGTAATTAAGTTTATCTCTGCTGGTGGTGGTGG 2432
QY 1452 GCCTGCTGCTCAACCCACTGCTGTACTCTCTTCAACCCCACTTCCGGGAGTACCT 1511
Db 2433 TCCTGCACTGTCTCAATCCCTTCTCTACATCTTGTTCATCTCTCACTTTAAGGAGATCT 2492
QY 1512 TCGGCGGCT 1520
Db 2493 GGTGAGCCT 2501
```

RESULT 6

```
US-09-495-050A-214
; Sequence 214, Application US/09495050A
; Patent No. 6492505
; GENERAL INFORMATION:
; APPLICANT: Roopa, Reddy
; APPLICANT: Guegler, Karl, J.
; APPLICANT: Au-Young, Janice
; TITLE OF INVENTION: COMPOSITION FOR DETECTION OF GENES ENCODING MEMBRANE-ASSOCIATED P
; FILE REFERENCE: PA-0013 US
; CURRENT APPLICATION NUMBER: US/09/495,050A
; PRIORITY FILING DATE: 2000-01-31
; PRIOR APPLICATION NUMBER: 60/118,318
; PRIOR FILING DATE: February 1, 1999
; NUMBER OF SEQ ID NOS: 305
; SOFTWARE: PERL Program
; SEQ ID NO 214
; LENGTH: 2612
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No. 6492505 2470285CT1
US-09-495-050A-214
```

```
Query Match      10.2%; Score 194; DB 4; Length 2612;
Best Local Similarity 52.8%; Pred. No. 4e-35;
Matches 469; Conservative 0; Mismatches 410; Indels 9; Gaps 2;

QY 622 AGTCTCCAGTGTAGCCCTTACTCCAGGCCCTTCAAGCCCTGTGAGTACCTCTTTGAAGAC 681
Db 1226 ATTATCCATTGTACACCTTCAACAGGTGCTTTAAGCCCTGTGAATATTTACTGGGAAGC 1285
QY 682 TGGGCACTCCCGCTGGCGCTGTGGCCATCGTGTGCTCTCGTGTCTGCAATGGACTG 741
Db 1286 TGGATGATTCGCTTACTGTGTGTTTCTTCTGTTGTTGATTTATTTTCAACTGCTT 1345
QY 742 GTGCTGTGACCGTGTGCTGGCGGCGCTGCCCGCTGCCCGCTCAAGTTGTGTGTA 801
```

```
Db 1346 GTTATTTTAAACAACATTTGCATCTTG---TACATCACTGCCTTCCTCCAAATTTGTTATA 1402
Qy 802 GGTGGGATTCAGGGCCGCAACACCTTGACTGCAATTTCTGTGGCCTTCTAGCCTCAGTC 861
Db 1403 GCCTTGATTTCTGTCTAACTTATTCATGGAAATCTATCTAGGCATCCTAACTTTTCTT 1462
Qy 862 GATGCCCTGACCTTTTGGTCAGTTCTCTGAGTACGAGCGCGCTGGAGACGGGGCTAGGC 921
Db 1463 GATGCTGTCTCTGGGCAGATTCGTGAATTTGGCAATTTGGTGGAACTGGCAGTGGC 1522
Qy 922 TGCCGGGCACTGGCTTCTGCGAGTACTCTGGGTGGAGGCAATCGGTCTGCTGCTACT 981
Db 1523 TGCAAGTAGTGGGTTTCTTGCACTTTCTCTCAGAAAGTGCATATTTTATTAAATG 1582
Qy 982 CTGGCCGAGTGCAGTGCAGGCTCTCCGTCTCCTGTGTCTCGGGCTATGGGAAGTCCCC 1041
Db 1583 CTAGCAACTGTGCAAGAAGCTTATCTGCAAAAGATATAATGAAATAATGGGAAGCAAT 1642
Qy 1042 TCCCTGGGCAGCGTTTCGAGCAGGGGTCTTAGGCTGCCTGGCACTGGCAGGGCTGGCCGCC 1101
Db 1643 CATCTCAACAGTTCGGGTTGCTGCCCCCTTTGGCTTTCCTAGGTGCTACAGTAGCAGGC 1702
Qy 1102 GCATGCCCTGGCCCTCAGTGGGAGAAATACGGGGCTCCCCACCTGCTGCTGCCCTACGCG 1161
Db 1703 TGTGTTCCCTTTTCCATAGAGGGGAATATCTGTCATCACCCCTTTGTTGGCAT----- 1757
Qy 1162 CCACCTGAGGTACGACGAGCCCTGGGCTTCCACGTGGCCCTGGTGGTATGATGAATCC 1221
Db 1758 -TTCCACAGGTGAACCCCAATCATTAGGATTCATGTAACTGTTAGTCTATTAACTCA 1816
Qy 1222 TTCTGTTTCTGCTGCTGGCGGCTTACATCAACTGTAATGTAAGTGTGACCTGCCGGGGC 1281
Db 1817 CTAGCATTTTATTAAGCGCGTTATCTACACTAGCTATACCTGCAACTTGGAAAAGAG 1876
Qy 1282 GACTTTGAGCGCGTGTGGGACTGCGCCATGTGAGGCACGTGGCTGCTCATCTCGCA 1341
Db 1877 GACCTCTCAGAAAACTCAAACTAGCATGATTAAGCATGTGCTGGCTAATCTTCACC 1936
Qy 1342 GACGGCTCCTTACTGTCGGTGGCTTCTCAGCTTGCCTTCATGCTGCTGCTGGCTCTTC 1401
Db 1937 AATTGCACTCTTTTCTGCGCTGTGGCGTTTTTTCATTTGACCAATGATCACTGCAATC 1996
Qy 1402 CTTGTACGCCCGGAGCGCTCAAGTCTGCTGCTGCTGGTGGTGGCTGCCCTGCCCTGCTG 1461
Db 1997 TCTATACGCCCGGAAATGAATGAGTCTGTACTCTGATATTTTTCATTTGCTGCTGCTG 2056
Qy 1462 CTCACCCACTGCTGTACTGCTCTTCAACCCCACTTCCGGGAGAC 1509
Db 2057 CTGAATCCAGTCTGTATGTTTCTTCAACCCAAAGTTTAAAGAAGAC 2104
```

RESULT 7

US-08-866-757-1
; Sequence 1, Application US/08866757
; Patent No. 5858716

GENERAL INFORMATION:

; APPLICANT: ELSHOURBAGY, NABIL A
; APPLICANT: LI, XIAOTONG
; APPLICANT: BERGSMAN, DEBK J
; TITLE OF INVENTION: NOVEL 7TM RECEPTOR (H2CAA71)

; NUMBER OF SEQUENCES: 4

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: RATNER & PRESTIA

; STREET: P.O. BOX 980

; CITY: VALLEY FORGE

; STATE: PA

; COUNTRY: USA

; ZIP: 19482

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Diskette

; COMPUTER: IBM Compatible

; OPERATING SYSTEM: DOS

; SOFTWARE: FastSeq for Windows Version 2.0

; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/866,757
; FILING DATE: 30-MAY-1997
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: PRESTIA, PAUL F
; REGISTRATION NUMBER: 23,031
; REFERENCE/DOCKET NUMBER: GH-70055
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 610-407-0700
; TELEFAX: 610-407-0701
; TELEX: 846169
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4203 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cdna
; US-08-866-757-1

Query Match 10.2%; Score 194; DB 2; Length 4203;

Best Local Similarity 52.8%; Pred. No. 4.6e-35;

Matches 469; Conservative 0; Mismatches 410; Indels 9; Gaps 2;

```
Qy 622 AGTGTCCAGTGTAGCCCTACTTCCAGGCCCCCTTCAAGCCCTGTGAGTACCTCTTTGAAGC 681
Db 1117 ATTATCCATTGTACACCTTCAACAGAGTGTCTTTAAGCCCTGTGAATATTTACTGGAGC 1176
Qy 682 TGGGGATTCGCCCTGGCCGCTGTGGCCCATCGTGTGCTCTCCGTCCTGCAATGGACTG 741
Db 1177 TGGATGATTCGCTCTACTGTGTGCTTCAATTTCTGTTGCTGCAATTTTCAACCTGCTT 1236
Qy 742 GTGCTGTGACCGGTTCGCTGGCGGCCCTGCCCCCTGCCCGGTCAAGTTTGTGGTA 801
Db 1237 GTTATTTTAAACAATTTGTCATCTTG---TACATCACTGCTTCTGTCMAATTTGTTATA 1293
Qy 802 GGTGCGATTCAGCGCCCAACACCTTGACTGGCAATTTCTGTGGCCTTCTAGCCTCAGTC 861
Db 1294 GGCTTGATTTCTGTGCTAACTTATTCATGGAATCTATCTAGGCATCTTAACTTTTCTT 1353
Qy 862 GATGCCCTGACCTTTGCTGAGTCTCTGAGTACGAGCCCGCTGGGAGACGGGCTAGGC 921
Db 1354 GATGCTGTGCTCTGGGCGAGATTCGCTGAAATTTGGCAATTTGGTGGGAAACTGGCAGTGC 1413
Qy 922 TGCCGGGCCACTGGCTTCTTCCGAGTACTTTGGGTGGGAGCATCGGTGCTGCTCACT 981
Db 1414 TGCAAGTAACTGGGTTTCTTGCAAGTTTCTCTCAGAAAGTGCATATTTTATTAAATG 1473
Qy 982 CTGGCCGAGTGCAGTGCAGCGTCTCCGTCTCTCTGTCTCCGGGCTTATGGGAAGTCCCC 1041
Db 1474 CTAGCAACTGTCGAAAGAAAGCTTATCTGCAAAAAGATATAATGAAATAATGGAGAGCAAT 1533
Qy 1042 TCCCTGGGAGCGTTTCGAGCAGGGGCTCTAGGTGCTGCTGAGCCTGGCAGGCTGGCCGCC 1101
Db 1534 CATCTCAACAGCTTCCGGGTGCTGCCCCCTTTTGGCTTTTCTGAGTGTCTACAGTAAACAGGC 1593
Qy 1102 GCACTGCCCTGGCCTCAGTGGGAGAAATACGGGCGCTCCCACTCTCCCTGCCCTACGCG 1161
Db 1594 TGTGTTTCCCTTTTCCATAGAGGGGAATATCTGTCATCACCCCTTTGTTTGGCAT----- 1648
Qy 1162 CCACCTGAGGCTCAGCAGCAGCCCTGGGGCTTCCCGTGGCCCTGGTGGTATGATGAATCC 1221
Db 1649 -TTCTACAGGTGAACCGCCATCATTAGGATTCACCTGAACGTTAGTGTCTATTAAACTCA 1707
Qy 1222 TTCTGTTTCTGCTGCGCGGTGCTCATCAAACTGCTGCTGACCTGCCCGGGGGC 1281
Db 1708 CTAGCATTTTTTATTAATGGCGGTTATCTACACTAAAGCTATATGCAACTTGGAAAAAGAG 1767
Qy 1282 GACTTTGAGCGCGTGTGGGACTGCGCCATGGTGAGGACGCTGGGCTGGCTCATCTTCGCA 1341
```

Db 1768 GACCTCTCAGAAACCTCAAAATCTAGCATGATTAAGCATGTGCGTTGGCTAAATCTTCACC 1827
Qy 1342 GACGGGCTCCTACTGTGCCGCTTCCCTCAGCTTCGCCTCCATGCTGGGCTCTTC 1401
Db 1828 AATTGCATCTTTTCTGCGCTGTGGGTTTTTTCATTTCACCATGATCACTCAATC 1887
Qy 1402 CTTGTACGCGCGGCGCGTCAAGTCTGTCTGTCTGTGGTGTGCTGCCCTGCTGCTGCG 1461
Db 1888 TCTATCAGCGCGGAAATATGAAGTCTGTACTCTGATATTTTTCATTGCTGCTGTC 1947
Qy 1462 CTCACCCACCTGCTGCTGCTCTTCAACCCCACTTCGCGGATGAC 1509
Db 1948 CTGAATCCAGTCTGTATGTTTTTCTCAACCCCAAGTTTAAAGAAGAC 1995

RESULT 8

US-09-153-593-1

; Sequence 1, Application US/09153593A

; Patent No. 6174994

; GENERAL INFORMATION:

; APPLICANT: ELSHOURBAGY, NABIL A

; APPLICANT: LI, XIAOTONG

; APPLICANT: BERGSMAN, DEBK J

; TITLE OF INVENTION: NOVEL 7TM RECEPTOR (H2CAA71)

; FILE REFERENCE: GH-70055-1

; CURRENT APPLICATION NUMBER: US/09/153,593A

; EARLIER FILING DATE: 1998-09-15

; EARLIER FILING DATE: 08/866,757

; NUMBER OF SEQ ID NOS: 4

; SOFTWARE: FastSeq for Windows Version 3.0

; SEQ ID NO 1

; LENGTH: 4203

; TYPE: DNA

; ORGANISM: HOMO SAPIENS

US-09-153-593-1

Query Match 10.2%; Score 194; DB 3; Length 4203;
Best Local Similarity 52.8%; Pred. No. 4.6e-35;
Matches 469; Conservative 0; Mismatches 410; Indels 9; Gaps 2;

Qy 622 AGTGTCCAGTGTAGCCCTACTCCAGGCCCTTCAAGCCCTGTGAGTACTCTTTGAAAGC 681
Db 1117 ATTATCCATTGTACACTTCAACAGGTGCTTTTAAGCCCTGTGAATATTTACTTGGGAGC 1176
Qy 682 TGGGSCATCCGCTGGCGGTGGGCCATCGTTGCTCTCGTCTCGTCAATGACGTG 741
Db 1177 TGGATGATTCGTCTTACTGTGTGGTTCAATTTCTTGGTTGCAATATTTTCAACCTGCTT 1236
Qy 742 GTGCTGTGACCGTGTTCGCTGGCGGCTGCCCCCTGCCCGGTCAAGTTTGTGGTA 801
Db 1237 GTTATTTTAAACAACTTGTGATCTTG---TACATCACTGCTTCTGCCAAATGTTTATA 1293
Qy 802 GGTGCGATTGAGGCGCAACACCTTGATGTCATTTCTGTGGCTTTCTAGCCTCAGTC 861
Db 1294 GCGTGTATTTCTGTCTAACTTATTTCATGGAATCTACTTGGCATCTTAACCTTTCTT 1353
Qy 862 GATGCCCTGACCTTTGGTCAGTCTCTGAGTACGAGCCGCTGGAGACCGGGCTAGC 921
Db 1354 GATGCTGTCTTGGGGCAGATTGCTGAAATTTGGCAATTTGGTGGAACTTGGCAGTGC 1413
Qy 922 TGCCGGGCACTTGGCTTCTGTCAGTACTTGGGTGGAGGATCGGTGCTGCTGCTCACT 981
Db 1414 TGCAAGTAACTGGGTTTCTGAGTTTCTTCTCAGAAAGTGCATATTTTATTATG 1473
Qy 982 CTGCGCGCAGTGCAGTGCAGCGTCTCCGTCTCTGTCTGTCCGGGCTATGGGAAGTCCCC 1041
Db 1474 CTAGCAACTGTGCAAGAAGCTTATCTGCAAAAGATATAATGAAATAATGGGAAGCAAT 1533
Qy 1042 TCCCTGGGCGAGGTTGAGAGAGGGGTCTTAGCTGCTGGCACTGGCAGGCGTGGCCGCC 1101
Db 1534 CATCTCAAAACAGTTCCGGGTTGCTGCCCTTTTGGCTTTTCTAGGTGCTACAGTAACAGGC 1593

RESULT 9

US-09-495-050A-220

; Sequence 220, Application US/09495050A

; Patent No. 6492505

; GENERAL INFORMATION:

; APPLICANT: Roopa, Reddy

; APPLICANT: Guegler, Karl, J.

; APPLICANT: Au-Young, Janice

; TITLE OF INVENTION: COMPOSITION FOR DETECTION OF GENES ENCODING MEMBRANE-ASSOCIATED P

; FILE REFERENCE: PA-0013 US

; CURRENT APPLICATION NUMBER: US/09/495,050A

; CURRENT FILING DATE: 2000-01-31

; PRIOR APPLICATION NUMBER: 60/118,318

; PRIOR FILING DATE: February 1, 1999

; NUMBER OF SEQ ID NOS: 305

; SOFTWARE: PERL Program

; SEQ ID NO 220

; LENGTH: 723

; TYPE: DNA

; ORGANISM: Homo sapiens

; NAME/KEY: misc feature

; OTHER INFORMATION: Incyte ID No. 6492505 2603450CT1

US-09-495-050A-220

Query Match 9.6%; Score 181.8; DB 4; Length 723;
Best Local Similarity 55.8%; Pred. No. 1.8e-32;
Matches 392; Conservative 0; Mismatches 302; Indels 9; Gaps 2;

Qy 695 TGGCGGTGTGGGCGCATCGTGTGCTCTCCGTGCTCTGCAATGGACTGGTGTGTCGACCG 754
Db 1 TTGAGTGTGACCATAGCAGTTCTGGCAGTTACTTGTAAATGCTTGGTGTGACTTCAACAG 60
Qy 755 TGTTTCGTGGGCGCTGCCCCCTGCCCGGTCAAGTTTGTGTAGTGTGATTTGCGAG 814
Db 61 TTTTC---AGATCCCTCTGTACATTTCCCCCAATTAACCTGTAATTTGGGTGTCATCGAG 117
Qy 815 GCGCAACACCTTGACTGGCAATTTCTGTGGCTTTCTAGCCTTCAGTCCATGCCCTGACCT 874
Db 118 CAGTGAACATGCTCAGCGGAGTCTCCAGTCCGCTGGCTGGTGTGGATGGTCACTT 177
Qy 875 TTGTCAGTTCTCTGAGTACGAGCCCGCTGGGAGAGCGGGCTAGGCTGCGGGCCTAGT 934

Db 74 CCGTATTCTACACTAAGCTATTAATGCAACTTGGAAAAAGAGGACCTCTCAGAAAACTCAC 133
Qy 1301 ACTGCGCCATGCTGAGGCACGCTGGCTGGCTCATCTTGCAGACGGGCTCCTCTACTGTC 1360
Db 134 AATCTAGCATGATTAAGCATGTCGCTGGCTAATCTTACCNAATGTCATCTTTTCTGCC 193
Qy 1361 CCGTGGCCTTCTCAGCTTGGCTCCATGCTGGGCTCTTCCCTGTACGCCCCGAGGCGG 1420
Db 194 CTGTGGCGTTTTTTTCATTTGACCACTGATCACTGCAATCTCTATCAGCCCCGAAATAA 253
Qy 1421 TCAAGTCTGCTCTGCTGGTGGCTGCCCCCTGCCTGCCTCAACCACTGCTGTACC 1480
Db 254 TGAAGTCTGTTACTCTGATATTTTTCATTTGCCCTGGCTTGCCTGAATCCAGTCTGTATG 313
Qy 1481 TGCTCTTCAACCCCACTTCCGGGATGAC 1509
Db 314 TTTTCTTCAACCAAGTTTAAAGAGGAC 342

RESULT 12
US-09-153-593-3
; Sequence 3, Application US/09153593A
; Patent No. 6174994
; GENERAL INFORMATION:
; APPLICANT: ELSHOURBAGY, NABIL A
; APPLICANT: LI, XIAOTONG
; APPLICANT: BERGSMAN, DEBK J
; TITLE OF INVENTION: NOVEL 7TM RECEPTOR (H2CAA71)
; CURRENT APPLICATION NUMBER: US/09/153,593A
; CURRENT FILING DATE: 1998-09-15
; EARLIER APPLICATION NUMBER: 08/866,757
; EARLIER FILING DATE: 1997-05-30
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 3
; LENGTH: 473
; TYPE: DNA
; ORGANISM: HOMO SAPIENS
; FEATURE:
; NAME/KEY: UNSURE
; LOCATION: (397) (400) (406) (432) (459)
US-09-153-593-3

Query Match 5.4%; Score 101.8; DB 3; Length 473;
Best Local Similarity 56.8%; Pred. No. 4.7e-14;
Matches 187; Conservative 0; Mismatches 142; Indels 0; Gaps 0;
Qy 1181 CAGCCCTGGGCTTACCGTGGCCCTGGTGATGAATCACTCTTCTGTTTCTGCTGGTGG 1240
Db 14 CATCATTAGGATTCACTGTAACTGCTGCTATTAATCACTAGCAATTTTATTAAATGG 73
Qy 1241 CCGTGGCTTACATCAACTGATCTGTGACCTGCCGGGGGACCTTGAAGCCGTGGG 1300
Db 74 CCGTATTCTACACTAAGCTATTAATGCAACTTGGAAAAAGAGGACCTCTCAGAAAACTCAC 133
Qy 1301 ACTGCGCCATGCTGAGGCACGCTGGCTCATCTTGCAGACGGGCTCCTCTACTGTC 1360
Db 134 AATCTAGCATGATTAAGCATGTCGCTGGCTAATCTTACCNAATGTCATCTTTTCTGCC 193
Qy 1361 CCGTGGCCTTCTCAGCTTGGCTCCATGCTGGGCTCTTCCCTGTACGCCCCGAGGCGG 1420
Db 194 CTGTGGCGTTTTTTTCATTTGACCACTGATCACTGCAATCTCTATCAGCCCCGAAATAA 253
Qy 1421 TCAAGTCTGCTGCTGGTGGCTGCCCCCTGCCTGCCTCAACCACTGCTGTACC 1480
Db 254 TGAAGTCTGTTACTCTGATATTTTTCATTTGCCCTGGCTTGCCTGAATCCAGTCTGTATG 313
Qy 1481 TGCTCTTCAACCCCACTTCCGGGATGAC 1509
Db 314 TTTTCTTCAACCAAGTTTAAAGAGGAC 342

RESULT 13
US-09-482-273-101
; Sequence 101, Application US/09482273
; Patent No. 6534631
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: 71 Human Secreted Proteins
; FILE REFERENCE: P2030P1
; CURRENT APPLICATION NUMBER: US/09/482,273
; CURRENT FILING DATE: 2000-01-13
; EARLIER APPLICATION NUMBER: PCT/US99/15849
; EARLIER FILING DATE: 1999-07-14
; EARLIER APPLICATION NUMBER: 60/092,921
; EARLIER FILING DATE: 1998-07-15
; EARLIER APPLICATION NUMBER: 60/092,922
; EARLIER FILING DATE: 1998-07-15
; EARLIER APPLICATION NUMBER: 60/092,956
; EARLIER FILING DATE: 1998-07-15
; NUMBER OF SEQ ID NOS: 267
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 101
; LENGTH: 2709
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-482-273-101

Query Match 5.0%; Score 94.6; DB 4; Length 2709;
Best Local Similarity 56.5%; Pred. No. 3.5e-12;
Matches 195; Conservative 0; Mismatches 149; Indels 1; Gaps 1;
Qy 1165 CCGAGGCTCAGCCAGCAGCCCTGGGCTTACCGTGGCCCTGGTGATGATGAATCCTTC 1224
Db 13 CCGTACAGGTGAACCGCCATCATTAGGATTCCTGTAACTGCTATTAACTCACTA 72
Qy 1225 TGTTCCTGGTGGCGGCTGCTACATCAACTGATCTGCTGACCTGCCGGGGCCAC 1284
Db 73 GCATTTTATTAAATGGCGCTTATCTACACTAAGCTATATCTCAACTTGGAAAAAGAGGAC 132
Qy 1285 TTGAGGCGGTGGGACTGGCCATGGTGAAGGACAGTGGGCTGCTCATCTTCGACAC 1344
Db 133 CTCTCAGAAAACTCACATCTAGCATGATTAAGCATGCTGGTTGGCTAATCTTACCAT 192
Qy 1345 GGGCTCCTCTACTGTCCGTCGCTTCTCAGCTTGGCTCCATGCTGGGCTCTTCCCT 1404
Db 193 TGCATCTTTTCTGCCCTGTGGCGTTTTTTTCAITTTGCACCATGATCACTGCAATCTCT 252
Qy 1405 GTCAGGCCGAGGCGGTCAAGTCTGCTGCTGGTGGTGGCTGCCCTGCCCTGCCCTC 1464
Db 253 ATCAGCCCCGAAATAATGAAGTCTGTTACTCTGAT-ATTTTTTCCATGCCCTGCTG 311
Qy 1465 AACCCACTGCTGACCTGCTCTTCAACCCCACTTCCGGGATGAC 1509
Db 312 AATCAGTCTCTGATGTTTCTTCAACCAAGCTTTAAAGAGGAC 356

RESULT 14
US-09-589-510-7
; Sequence 7, Application US/09589510
; Patent No. 6706449
; GENERAL INFORMATION:
; APPLICANT: Mahajan, Pramod B.
; TITLE OF INVENTION: Maize Orthologues of Bacterial RuvB:
; FILE REFERENCE: 1121
; CURRENT APPLICATION NUMBER: US/09/589,510
; CURRENT FILING DATE: 2000-06-07
; PRIOR APPLICATION NUMBER: US 60/144,112
; PRIOR FILING DATE: 1999-07-16
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 7
; LENGTH: 1898
; TYPE: DNA

```
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (166)...(1533)
US-09-589-510-7

Query Match      4.2%; Score 79.8; DB 4; Length 1898;
Best Local Similarity 80.9%; Pred. No. 8.2e-09;
Matches 93; Conservative 0; Mismatches 22; Indels 0; Gaps 0;

Qy 1 AATACGACTCTACTATAGGAAAGCTGGTACCGCTGCAGGTACCGGTCCGGAATTCCTCGG 60
Db 16 AATACGACTCTACTATAGGAAAGCTGGTACCGCTGCAGGTACCGGTCCGGAATTCCTCGG 75

Qy 61 TCGACCCACGCGTCCGTGGAGCGGAGCCAGGGTCTGAGCCTGCCGGCTCATCCAG 115
Db 76 TCGACCCACGCGTCCGTGGAGCGGCGCGCGCGGAGAGAGCAGCTCTAG 130
```

```
RESULT 15
US-08-633-682-1
; Sequence 1, Application US/08633682
; Patent No. 5840544
; GENERAL INFORMATION:
; APPLICANT: Hawkins, Phillip R.
; APPLICANT: Bandman, Olga
; APPLICANT: Murry, Lynn E.
; TITLE OF INVENTION: NOVEL RANTES HOMOLOG FROM PROSTATE
; NUMBER OF SEQUENCES: 5
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Drive
; CITY: Palo Alto
; STATE: CA
; COUNTRY: U.S.
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq Version 1.5
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/633,682
; FILING DATE: Filed Herewith
; ATTORNEY/AGENT INFORMATION:
; NAME: Luther, Barbara J
; REGISTRATION NUMBER: 33,954
; REFERENCE/DOCKET NUMBER: PF-0063 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-855-0555
; TELEFAX: 415-852-0195
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 570 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; IMMEDIATE SOURCE:
; LIBRARY: Prostrate
; CLONE: 836820
US-08-633-682-1
```

```
Query Match      4.1%; Score 78.8; DB 2; Length 570;
Best Local Similarity 73.2%; Pred. No. 1e-08;
Matches 101; Conservative 0; Mismatches 37; Indels 0; Gaps 0;

Qy 1 AATACGACTCTACTATAGGAAAGCTGGTACCGCTGCAGGTACCGGTCCGGAATTCCTCGG 60
Db 11 AATACGACTCTACTATAGGAAAGCTGGTACCGCTGCAGGTACCGGTCCGGAATTCCTCGG 70

Qy 61 TCGACCCACGCGTCCGTGGAGCGGAGCCAGGGTCTGAGCCTGCCGGCTCATCCAGCTCT 120
```

```
Db 71 TCGACCCACGCGTCCGAGGGCGCTGATTTGCAGCATCATGATGGCGCTCTCTCTGGCCTCT 130
Qy 121 CTTGCTGCCCTAGCGGC 138
Db 131 GCTGTGCTCTCTGGCCTCC 148

Search completed: April 14, 2005, 18:45:42
Job time : 223.132 secs
```

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 14, 2005, 12:03:58 ; Search time 720.624 Seconds
(without alignments)
15987.096 Million cell updates/sec

Title: US-10-664-667-6
Perfect score: 1899
Sequence: 1 aatacgaactactatagga.....tggccttggcttcacagtg 1899

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 5622541 seqs, 303335566 residues
Total number of hits satisfying chosen parameters: 11245082

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications NA.*
1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq.*
2: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq.*
3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq.*
4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq.*
5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq.*
6: /cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq.*
7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq.*
8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq.*
9: /cgn2_6/ptodata/1/pubpna/US09A_PUBCOMB.seq.*
10: /cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq.*
11: /cgn2_6/ptodata/1/pubpna/US09C_PUBCOMB.seq.*
12: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq.*
13: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq.*
14: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq.*
15: /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq.*
16: /cgn2_6/ptodata/1/pubpna/US10D_PUBCOMB.seq.*
17: /cgn2_6/ptodata/1/pubpna/US10E_PUBCOMB.seq.*
18: /cgn2_6/ptodata/1/pubpna/US10F_PUBCOMB.seq.*
19: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq.*
20: /cgn2_6/ptodata/1/pubpna/US11_NEW_PUB.seq.*
21: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq.*
22: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1898	99.9	1899	10	US-09-851-595-6
2	1898	99.9	1899	17	US-10-664-667-6
3	1898	99.9	2486	10	US-09-851-595-4
4	1898	99.9	2486	17	US-10-664-667-4
5	1754.4	92.4	2208	10	US-09-851-595-9
6	1754.4	92.4	2208	17	US-10-664-667-9
7	1754.4	92.4	2711	10	US-09-851-595-7
8	1754.4	92.4	2711	17	US-10-664-667-7
9	1754.4	92.4	2901	10	US-09-851-595-12
10	1754.4	92.4	2901	17	US-10-664-667-12
11	1754.4	92.4	3325	17	US-10-331-496A-52
					Sequence 6, Appli
					Sequence 6, Appli
					Sequence 4, Appli
					Sequence 9, Appli
					Sequence 9, Appli
					Sequence 7, Appli
					Sequence 7, Appli
					Sequence 12, Appl
					Sequence 12, Appl
					Sequence 52, Appl

12	1754.4	92.4	3438	14	US-10-176-847-89	Sequence 89, Appl
13	1754.4	92.4	3443	17	US-10-331-496A-94	Sequence 94, Appl
14	1754.4	92.4	3492	10	US-09-851-595-10	Sequence 10, Appl
15	1754.4	92.4	3492	17	US-10-664-667-10	Sequence 10, Appl
16	1754.4	92.4	3492	18	US-10-737-450-31	Sequence 31, Appl
17	1752.8	92.3	2988	17	US-10-302-172-723	Sequence 723, App
18	1752.8	92.3	3119	15	US-10-225-567A-580	Sequence 580, App
19	1751.2	92.2	3429	18	US-10-398-036-26	Sequence 26, Appl
20	1749.6	92.1	2487	14	US-10-270-336-1	Sequence 1, Appli
21	1748	92.0	3273	17	US-10-295-027-1119	Sequence 1119, Ap
22	1748	92.0	3273	18	US-10-783-528-43	Sequence 43, Appl
23	1746.4	92.0	3349	14	US-10-270-336-4	Sequence 4, Appli
24	1740.8	91.7	3381	10	US-09-970-944-5	Sequence 5, Appli
25	1336	70.4	2901	10	US-09-851-595-3	Sequence 3, Appli
26	1336	70.4	2901	17	US-10-664-667-3	Sequence 3, Appli
27	1336	70.4	3637	10	US-09-851-595-1	Sequence 1, Appli
28	1336	70.4	3637	17	US-10-664-667-1	Sequence 1, Appli
29	1254.8	66.1	119596	14	US-10-270-336-3	Sequence 3, Appli
30	723.2	38.1	1173	17	US-10-343-650A-65	Sequence 65, Appl
31	493.6	26.0	1078	17	US-10-264-049-395	Sequence 395, App
32	439.4	23.1	453	9	US-09-876-889-240	Sequence 240, App
33	416.2	21.9	422	10	US-09-918-995-34626	Sequence 34626, A
34	356.2	18.8	2082	19	US-10-851-470-3	Sequence 3, Appli
35	355.6	18.7	2651	17	US-10-295-027-483	Sequence 483, App
36	355.6	18.7	2651	17	US-10-173-999-27	Sequence 27, Appl
37	355.6	18.7	2724	15	US-10-251-385-263	Sequence 263, App
38	355.6	18.7	2724	15	US-10-225-567A-421	Sequence 421, App
39	355.6	18.7	2724	17	US-10-174-456-4	Sequence 4, Appli
40	355.6	18.7	2724	18	US-10-751-736-21	Sequence 21, Appl
41	355.6	18.7	2880	17	US-10-295-027-1114	Sequence 1114, Ap
42	355.6	18.7	2880	19	US-10-482-029-157	Sequence 157, App
43	355.6	18.7	2880	19	US-10-651-237-48	Sequence 48, Appl
44	355.6	18.7	2880	19	US-10-782-413-48	Sequence 48, Appl
45	355.6	18.7	2880	19		

ALIGNMENTS

RESULT 1
US-09-851-595-6
; Sequence 6, Application US/09851595
; Publication No. US20030166047A1
; GENERAL INFORMATION:
; APPLICANT: Gu, Wei
; TITLE OF INVENTION: NOVEL G-PROTEIN COUPLED RECEPTORS AND USES THEREFOR
; FILE REFERENCE: MNI-080CP
; CURRENT APPLICATION NUMBER: US/09/851,595
; CURRENT FILING DATE: 2000-05-08
; PRIOR APPLICATION NUMBER: 09/556,588
; PRIOR FILING DATE: 2000-05-08
; PRIOR APPLICATION NUMBER: 60/132,896
; PRIOR FILING DATE: 1999-05-06
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 6
; LENGTH: 1899
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(1899)
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (171)
; OTHER INFORMATION: n = any nucleotide
US-09-851-595-6

Query Match 99.9%; Score 1898; DB 10; Length 1899;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1899; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 AATACGACTACTATAGGGAAGTGGTACGCTCGAGGTACCGTCCGGAATCCCGG 60

Db 1 AATACGACTCATATAGGAAAGCTGGTACGCTCGAGGTACCGGTCCGGAATTCGCCGG 60
Qy 61 TCGACCCAGCGGTCCGTGGAGGGAGCCAGGGTCTGAGCCCTGCGGCTCATCCAGCCTCT 120
Db 61 TCGACCCAGCGGTCCGTGGAGGGAGCCAGGGTCTGAGCCCTGCGGCTCATCCAGCCTCT 120
Qy 121 CTTGCTGCCCTAGCGGCTCCAAACAACCGCATCTGGGAAATTGGAGCTNGACACCTTC 180
Db 121 CTTGCTGCCCTAGCGGCTCCAAACAACCGCATCTGGGAAATTGGAGCTNGACACCTTC 180
Qy 181 AGCCAGCTGAGTCCCTCAAGCCCTGGATCTTAGCTGGAAACGCCATCCGGTCCATCCAC 240
Db 181 AGCCAGCTGAGTCCCTCAAGCCCTGGATCTTAGCTGGAAACGCCATCCGGTCCATCCAC 240
Qy 241 CTTGAGGCTCTTCCACCTGCACTCCCTGGTCAAGCTGGAGCTGAGACAGAACACAGCTG 300
Db 241 CTTGAGGCTCTTCCACCTGCACTCCCTGGTCAAGCTGGAGCTGAGACAGAACACAGCTG 300
Qy 301 ACCACACTGCCCTGGCTGGACTTGGGGCTTGTATGCACTGAAAGCTCAAAAGGGAACCTT 360
Db 301 ACCACACTGCCCTGGCTGGACTTGGGGCTTGTATGCACTGAAAGCTCAAAAGGGAACCTT 360
Qy 361 GCTCTCTCCAGGCTCTTCCAAAGGACAGTCTTCCAAAACTGAGGATCCTTGGAGTGCCT 420
Db 361 GCTCTCTCCAGGCTCTTCCAAAGGACAGTCTTCCAAAACTGAGGATCCTTGGAGTGCCT 420
Qy 421 TATGCTACAGTGTCTCCCTATGGGATGTGTGCGAGCTTCTTCAAGGCTCTTGGGAG 480
Db 421 TATGCTACAGTGTCTCCCTATGGGATGTGTGCGAGCTTCTTCAAGGCTCTTGGGAG 480
Qy 481 TGGGAGGCTGAAGACCTTCACTTGTATGATGAGAGTCTTCAAAAAGGCCCTTGGGCTC 540
Db 481 TGGGAGGCTGAAGACCTTCACTTGTATGATGAGAGTCTTCAAAAAGGCCCTTGGGCTC 540
Qy 541 CTTGCCAGACAAGCAGAGAACCATATGACAGGACCTTGGATGAGCTCCAGCTGGAGATG 600
Db 541 CTTGCCAGACAAGCAGAGAACCATATGACAGGACCTTGGATGAGCTCCAGCTGGAGATG 600
Qy 601 GAGACTCAAAAGCAGACCCAGTGTCCAGTGTAGCTTCAAGGCTTCAAGGCC 660
Db 601 GAGACTCAAAAGCAGACCCAGTGTCCAGTGTAGCTTCAAGGCTTCAAGGCC 660
Qy 661 TGTGAGTACCTTTTGAAGCTGGGGCATCGCTGGCCGTGGGCTCATCGTGTGCTC 720
Db 661 TGTGAGTACCTTTTGAAGCTGGGGCATCGCTGGCCGTGGGCTCATCGTGTGCTC 720
Qy 721 TCCGTGCTCTGCAATGGACTGGTGTGCTGACCGTGTGCTGGCGGCTTGCCTCCCTG 780
Db 721 TCCGTGCTCTGCAATGGACTGGTGTGCTGACCGTGTGCTGGCGGCTTGCCTCCCTG 780
Qy 781 CCCCCTGCAAGTTTGTGAGTGGATGCGATTCGAGGCGGCAACACCTTGAATGGCAATTC 840
Db 781 CCCCCTGCAAGTTTGTGAGTGGATGCGATTCGAGGCGGCAACACCTTGAATGGCAATTC 840
Qy 841 TGTGGCTCTTAGCTCAGTGCAGTGCCTGACCTTGTGTCAGTTCCTGAGTACGAGGCC 900
Db 841 TGTGGCTCTTAGCTCAGTGCAGTGCCTGACCTTGTGTCAGTTCCTGAGTACGAGGCC 900
Qy 901 CGCTGGAGAGCGGGCTAGGCTGCGGGGCCACTGGCTTCTGGCAGTACTTGGGTGCGAG 960
Db 901 CGCTGGAGAGCGGGCTAGGCTGCGGGGCCACTGGCTTCTGGCAGTACTTGGGTGCGAG 960
Qy 961 GCATCGGTGCTGCTCACTCTGCGCGCAGTGCAGTGCAGCGTCTCGTCTCTGTGTC 1020
Db 961 GCATCGGTGCTGCTCACTCTGCGCGCAGTGCAGTGCAGCGTCTCGTCTCTGTGTC 1020
Qy 1021 CGGSCCTATGGGAAGTCCCTCCTCGGAGCGTTCGAGCAGGGCTTCAAGGCTGCTG 1080
Db 1021 CGGSCCTATGGGAAGTCCCTCCTCGGAGCGTTCGAGCAGGGCTTCAAGGCTGCTG 1080
Qy 1081 GCACTGCGAGGCTGGCGCGCAGCTGCCCCCTCGCTCAGTGGGAGAAATACGGGCTCC 1140

Db 1081 GCACTGCGAGGCTGGCGCGCAGCTGCCCCCTGCTCAGTGGGAGAAATACGGGCTCC 1140
Qy 1141 CCACCTGCTCCCTAGCGCCACCTGAGGGTCAAGCAGCAGCCCTGGGCTTCAACGCTG 1200
Db 1141 CCACCTGCTCCCTAGCGCCACCTGAGGGTCAAGCAGCAGCCCTGGGCTTCAACGCTG 1200
Qy 1201 GCCCTGCTGATGATAAATCTCTTCTGTTCTGCTGCTGGCCGCTGCTATCAAACTG 1260
Db 1201 GCCCTGCTGATGATAAATCTCTTCTGTTCTGCTGCTGGCCGCTGCTATCAAACTG 1260
Qy 1261 TACTGTCACTGCGCGGGCGACTTTGAGCCCGTGTGGAGCTGCGGCATGGTGAAGCAC 1320
Db 1261 TACTGTCACTGCGCGGGCGACTTTGAGCCCGTGTGGAGCTGCGGCATGGTGAAGCAC 1320
Qy 1321 GTGGCTGGCTCATCTTTGCGAGAGGGCTCCTTACTGTGCTGCTGCTTCTCAGCTTC 1380
Db 1321 GTGGCTGGCTCATCTTTGCGAGAGGGCTCCTTACTGTGCTGCTGCTTCTCAGCTTC 1380
Qy 1381 GCCTCCATGTGGGCTCTTCCCTGTCAAGCCCGAGGCGCTCAAGTCTGTCTGCTGCTG 1440
Db 1381 GCCTCCATGTGGGCTCTTCCCTGTCAAGCCCGAGGCGCTCAAGTCTGTCTGCTGCTG 1440
Qy 1441 GTGCTGCCCTGCTGCTGCTCAACCCACTGCTGTACTGCTTCAACCCCACTTC 1500
Db 1441 GTGCTGCCCTGCTGCTGCTCAACCCACTGCTGTACTGCTTCAACCCCACTTC 1500
Qy 1501 CGGATGACCTTCCGCGGCTTCCGCCCCGCGCAGGGGACTCAGGGCCCCCTAGCTATGCT 1560
Db 1501 CGGATGACCTTCCGCGGCTTCCGCCCCGCGCAGGGGACTCAGGGCCCCCTAGCTATGCT 1560
Qy 1561 GCGGCGCGGAGCTGGAGAGAGCTCTGTGATTTCTACCCAGGCGCTGTAGCTTCTCT 1620
Db 1561 GCGGCGCGGAGCTGGAGAGAGCTCTGTGATTTCTACCCAGGCGCTGTAGCTTCTCT 1620
Qy 1621 GATGTGATCTCAATCTTGGAGCTTCTGAAGCTTCTGAGCTGGGCGGCCCTAGCTAT 1680
Db 1621 GATGTGATCTCAATCTTGGAGCTTCTGAAAGCTTCTGAAAGCTGGGCGGCCCTAT 1680
Qy 1681 GGCTTCCCTCAGTGACCTCATCTCTGTGTCAGCAGCAGGGGCCCCCAGGCTGAGGGC 1740
Db 1681 GGCTTCCCTCAGTGACCTCATCTCTGTGTCAGCAGCAGGGGCCCCCAGGCTGAGGGC 1740
Qy 1741 AGCCATTTGTAGAGCCAGAGGGGAACCACTTTGGGAACCCCAACCTCCATCGATGGA 1800
Db 1741 AGCCATTTGTAGAGCCAGAGGGGAACCACTTTGGGAACCCCAACCTCCATCGATGGA 1800
Qy 1801 GAACTGCTGTGAGGCGAGAGGATCTACGCCAGCAGGTGAGGCTTGTGAGGGGTGGC 1860
Db 1801 GAACTGCTGTGAGGCGAGAGGATCTACGCCAGCAGGTGAGGCTTGTGAGGGGTGGC 1860
Qy 1861 GGCTTTGAGCCCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCT 1899
Db 1861 GGCTTTGAGCCCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCT 1899

RESULT 2

US-10-664-667-6

; Sequence 6, Application US/10664667

; Publication No. US20040058377A1

; GENERAL INFORMATION:

; APPLICANT: Gu, Wei

; TITLE OF INVENTION: NOVEL G-PROTEIN COUPLED RECEPTORS AND USES THEREFOR

; FILE REFERENCE: MNI-080CP

; CURRENT APPLICATION NUMBER: US/10/664,667

; CURRENT FILING DATE: 2003-09-18

; PRIOR APPLICATION NUMBER: US/09/851,595

; PRIOR FILING DATE: 2000-05-08

; PRIOR APPLICATION NUMBER: 09/556,588

; PRIOR FILING DATE: 2000-05-08

; PRIOR APPLICATION NUMBER: 60/132,896

; PRIOR FILING DATE: 1999-05-06

; NUMBER OF SEQ ID NOS: 12

; SOFTWARE: PatentIn Ver. 2.0

; SEQ ID NO 6									
; LENGTH: 1899									
; TYPE: DNA									
; ORGANISM: Homo sapiens									
; FEATURE:									
; NAME/KEY: CDS									
; LOCATION: (1)..(1899)									
; FEATURE:									
; NAME/KEY: misc feature									
; LOCATION: (171)									
; OTHER INFORMATION: n = any nucleotide									
US-10-664-567-6									
Query March 1999; Score 1898; DB 17; Length 1899;									
Best Local Similarity 100.0%; Pred. No. 0;									
Matches 1899; Conservative 0; Mismatches 0; Indels 0; Gaps 0									
Qy	1	AATACGACTCACTATAGGGAAGCTGGTACGCTGCAGGTACCGGTCCGGTAATTC	CGGG	60					
Db	1	AATACGACTCACTATAGGGAAGCTGGTACGCTGCAGGTACCGGTCCGGTAATTC	CGGG	60					
Qy	61	TCGACCCACGCTCCGTGGAGCGGAGGCTGTGAGCTGCCGGCTCATCCAGCCTCT	120						
Db	61	TCGACCCACGCTCCGTGGAGCGGAGGCTGTGAGCTGCCGGCTCATCCAGCCTCT	120						
Qy	121	CTTGCTGCCCTAGGGGCTCCAAACAACCGGATCTGGAAATTTGGAGCTTGACCTTC	180						
Db	121	CTTGCTGCCCTAGGGGCTCCAAACAACCGGATCTGGAAATTTGGAGCTTGACCTTC	180						
Qy	181	AGCCAGCTGAGCTCCCTGCAAGCCCTGGATCTTAGCTGGAAGCCCATCCGGTCCA	240						
Db	181	AGCCAGCTGAGCTCCCTGCAAGCCCTGGATCTTAGCTGGAAGCCCATCCGGTCCA	240						
Qy	241	CCTGAGGCTTCTCCACCTGCACCTCCTCTGGTCAAGCTGGACCTGACAGAAC	300						
Db	241	CCTGAGGCTTCTCCACCTGCACCTCCTCTGGTCAAGCTGGACCTGACAGAAC	300						
Qy	301	ACACACTGCCCTGGCTGGACTTGGGGCTTGATGCACTGAACTCAAAGGNA	360						
Db	301	ACACACTGCCCTGGCTGGACTTGGGGCTTGATGCACTGAACTCAAAGGNA	360						
Qy	361	GCTCTCTCCAGGCTTCTCCAAAGACAGTGTTC	420						
Db	361	GCTCTCTCCAGGCTTCTCCAAAGACAGTGTTC	420						
Qy	421	TATGCCTACAGTGTGTCTTATGGATGTGTGCCAGCTTCTTCAAGGCTCTGGG	480						
Db	421	TATGCCTACAGTGTGTCTTATGGATGTGTGCCAGCTTCTTCAAGGCTCTGGG	480						
Qy	481	TGGAGGCTGAAGACCTTCACCTTGATGATGAGGATCTTCAAAAGGCCCTTGG	540						
Db	481	TGGAGGCTGAAGACCTTCACCTTGATGATGAGGATCTTCAAAAGGCCCTTGG	540						
Qy	541	CTTCCACAGACAAGCAGAGAACCACTATACACAGACCTGGATGAGCTCCAG	600						
Db	541	CTTCCACAGACAAGCAGAGAACCACTATACACAGACCTGGATGAGCTCCAG	600						
Qy	601	GAGACTCAAAAGCACACCCCAAGTGTCCAGTGTAGCCCTATCTCCAGGCCCTT	660						
Db	601	GAGACTCAAAAGCACACCCCAAGTGTCCAGTGTAGCCCTATCTCCAGGCCCTT	660						
Qy	661	TGTGAGTACCTCTTTGAAAGCTGGGGCATCCGCTTGGCCGCTGTTGGTTC	720						
Db	661	TGTGAGTACCTCTTTGAAAGCTGGGGCATCCGCTTGGCCGCTGTTGGTTC	720						
Qy	721	TCCGTGTCTGCAATGGACTGGTGTCTGTGACCGTGTTTCGTGGGGGCTTGC	780						
Db	721	TCCGTGTCTGCAATGGACTGGTGTCTGTGACCGTGTTTCGTGGGGGCTTGC	780						
Qy	781	CCCCCGGTCAAGTTTGGTATAGTGCAGATTCGAGCGCCAAACCTTGACTGG	840						
Db	781	CCCCCGGTCAAGTTTGGTATAGTGCAGATTCGAGCGCCAAACCTTGACTGG	840						

<p>RESULT 3</p> <p>US-09-851-595-4</p> <p>; Sequence 4, Application US/09851595</p> <p>; Publication No. US20030166047A1</p> <p>; GENERAL INFORMATION:</p> <p>; APPLICANT: Gu, Wei</p> <p>; TITLE OF INVENTION: NOVEL G-PROTEIN COUPLED RECEPTORS AND USES THEREFOR</p> <p>; FILE REFERENCE: MNI-080CP</p> <p>; CURRENT APPLICATION NUMBER: US/09/851,595</p> <p>; CURRENT FILING DATE: 2000-05-08</p> <p>; PRIOR APPLICATION NUMBER: 09/556,588</p> <p>; PRIOR FILING DATE: 2000-05-08</p> <p>; PRIOR APPLICATION NUMBER: 60/132,896</p> <p>; PRIOR FILING DATE: 1999-05-06</p> <p>; NUMBER OF SEQ ID NOS: 12</p> <p>; SOFTWARE: PatentIn ver..2.0</p> <p>; SEQ ID NO 4</p> <p>; LENGTH: 2486</p> <p>; TYPE: DNA</p> <p>; ORGANISM: Homo sapiens</p> <p>; FEATURE:</p> <p>; NAME/KEY: CDS</p> <p>; LOCATION: (2)..(1900)</p> <p>; FEATURE:</p> <p>; NAME/KEY: misc feature</p> <p>; LOCATION: (172)</p> <p>; OTHER INFORMATION: n = any nucleotide</p> <p>US-09-851-595-4</p>									
Query Match	99.9%	Score 1898;	DB 10;	Length 2486;					
Best Local Similarity	100.0%	Pred. No. 0;							
Matches 1899;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;					
Qy	1	AATACGACTCTATAGGAAAGCTGTGACCCCTGCGAGGTACCGGTCCGGAAATTC	CGGG	60					
Db	2	AATACGACTCTATAGGAAAGCTGTGACCCCTGCGAGGTACCGGTCCGGAAATTC	CGGG	61					
Qy	61	TCGACCCACGCGTCCGTGGAGCGGAGCCAGGGTCTGAGCCTGCGGCTCATCCAGCCTCT	120						
Db	62	TCGACCCACGCGTCCGTGGAGCGGAGCCAGGGTCTGAGCCTGCGGCTCATCCAGCCTCT	121						
Qy	121	CTTGCTGCGCTTAGCGGCTCCAAACACACACCGCATCTGGGAAATTTGGAGCTNGACACCTTC	180						
Db	122	CTTGCTGCGCTTAGCGGCTCCAAACACACCGCATCTGGGAAATTTGGAGCTNGACACCTTC	181						
Qy	181	AGCCAGCTGAGCTCCCTGCAAGCCCTGGAATCTTAGCTGGAAGCGCATCCGGTCCATCCAC	240						
Db	182	AGCCAGCTGAGCTCCCTGCAAGCCCTGGAATCTTAGCTGGAAGCGCATCCGGTCCATCCAC	241						
Qy	241	CCTGAGGCTTCTCCACCCCTGCACTCCCTGCTCAAGCTGGAACCTGACAGCAACACGCTG	300						
Db	242	CCTGAGGCTTCTCCACCCCTGCACTCCCTGCTCAAGCTGGAACCTGGAACCAACGCTG	301						
Qy	301	ACCACACTGCGGCTTGGAGCTTGGGGCTTGATGATCTGAAAGCTCAAAAGGGAACCTT	360						
Db	302	ACCACACTGCGGCTTGGAGCTTGGGGCTTGATGATCTGAAAGCTCAAAAGGGAACCTT	361						
Qy	361	GCTCTCTCCAGGCTTCTCCAGGACAGTTTCCAAAACAGTATGAGGATCCTGGAGGTGCT	420						
Db	362	GCTCTCTCCAGGCTTCTCCAGGACAGTTTCCAAAACAGTATGAGGATCCTGGAGGTGCT	421						
Qy	421	TATGCTTACCAAGTCTGCTCCCTATGGGATGTGTGCCAGCTTCTTCAAGGCTCTGGGCAG	480						
Db	422	TATGCTTACCAAGTCTGCTCCCTATGGGATGTGTGCCAGCTTCTTCAAGGCTCTGGGCAG	481						
Qy	481	TGGAGGCTGAAGACCTTACCTTGAATGATGAGGAGTCTTCAAAAGGCTCTGGGCCTC	540						
Db	482	TGGAGGCTGAAGACCTTACCTTGAATGATGAGGAGTCTTCAAAAGGCTCTGGGCCTC	541						
Qy	541	CTTGCCACAGACAGAGAACCATATGACACAGGACCTGATGAGCTCCAGCTGGAGATG	600						
Db	542	CTTGCCACAGACAGAGAACCATATGACACAGGACCTGATGAGCTCCAGCTGGAGATG	601						

Qy	601	GAGGACTCAAAAGCCACACCCAGTGTCCAGTGTAGCCCTACTCCAGGCCCCCTTCAAGCCC	660
Db	602	GAGGACTCAAAAGCCACACCCAGTGTCCAGTGTAGCCCTACTCCAGGCCCCCTTCAAGCCC	661
Qy	661	TGTGAGTACCTCTTTTGAAGCTGGGGCATCCGCTGCGCGTGTGGGCGCATCGTGTGCTC	720
Db	662	TGTGAGTACCTCTTTTGAAGCTGGGGCATCCGCTGCGCGTGTGGGCGCATCGTGTGCTC	721
Qy	721	TCGGTGTCTTGAATGAGTGTGCTGCTGACCGTGTTCGCTGGCGGCGCTGCCCCCTG	780
Db	722	TCGGTGTCTTGAATGAGTGTGCTGCTGACCGTGTTCGCTGGCGGCGCTGCCCCCTG	781
Qy	781	CCCCCGTCAAGTTTGTGTAGTGCATTCGAGCGCCCAACACTTTGACTGGCATTTCC	840
Db	782	CCCCCGTCAAGTTTGTGTAGTGCATTCGAGCGCCCAACACTTTGACTGGCATTTCC	841
Qy	841	TGTGGCCTTCTAGCCTCAGTCGATGCCCTGACCTTTTGGTTCAGTTCTCTGAGTACGGAGCC	900
Db	842	TGTGGCCTTCTAGCCTCAGTCGATGCCCTGACCTTTTGGTTCAGTTCTCTGAGTACGGAGCC	901
Qy	901	CGCTGGGAGACGGGGCTAGGCTGCGGGCCACTGGGCTTCTTGGCAGTACTTGGGTCGAG	960
Db	902	CGCTGGGAGACGGGGCTAGGCTGCGGGCCACTGGGCTTCTTGGCAGTACTTGGGTCGAG	961
Qy	961	GCATCGGTGCTGTGCTCACTCTGGCGCAGTGAGTGACAGCGTCTCCGTCTCTCTGTGTC	1020
Db	962	GCATCGGTGCTGTGCTCACTCTGGCGCAGTGAGTGACAGCGTCTCCGTCTCTCTGTGTC	1021
Qy	1021	CGGGCCTATGGGAAGTCCCTCCCTGGGAGCGGTTTCGAGCAGGGGTCTTAGGCTGCGCTG	1080
Db	1022	CGGGCCTATGGGAAGTCCCTCCCTGGGAGCGGTTTCGAGCAGGGGTCTTAGGCTGCGCTG	1081
Qy	1081	GCATGGCAGGGCTGGCGCGCGCACTGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	1140
Db	1082	GCATGGCAGGGCTGGCGCGCGCACTGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	1141
Qy	1141	CCACTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	1200
Db	1142	CCACTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	1201
Qy	1201	GCCTCGTGTGATGAACTCCTTCTGTTTCTGTTTCTGTTTCTGTTTCTGTTTCTGTTTCTGTTT	1260
Db	1202	GCCTCGTGTGATGAACTCCTTCTGTTTCTGTTTCTGTTTCTGTTTCTGTTTCTGTTTCTGTTT	1261
Qy	1261	TACTGTGACCTGCGCGGGGCGAGTTTGGCCGCTGAGGAGCTGCGGCACTGTTGGTGGAGCAC	1320
Db	1262	TACTGTGACCTGCGCGGGGCGAGTTTGGAGGCGCTTGGAGGCTGCGGCACTGTTGGTGGAGCAC	1321
Qy	1321	GTGGCCTTGGCTCATCTTCGAGACGGGCTCCTCTACTGTCCCGTGGGCTTCTCCTCAGCTTC	1380
Db	1322	GTGGCCTTGGCTCATCTTCGAGACGGGCTCCTCTACTGTCCCGTGGGCTTCTCCTCAGCTTC	1381
Qy	1381	GCCTTCCATGCTGGGCTCTTTCCTGTGTCAGCCGAGGCGCTCAAGTCTGTCTGCTGGTG	1440
Db	1382	GCCTTCCATGCTGGGCTCTTTCCTGTGTCAGCCGAGGCGCTCAAGTCTGTCTGCTGGTG	1441
Qy	1441	GTGCTGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	1500
Db	1442	GTGCTGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	1501
Qy	1501	CGGGATGACCTTTCGGGGCTTTCGGGGCTTTCGGGGCTTTCGGGGCTTTCGGGGCTTTCGGGG	1560
Db	1502	CGGGATGACCTTTCGGGGCTTTCGGGGCTTTCGGGGCTTTCGGGGCTTTCGGGGCTTTCGGGG	1561
Qy	1561	GCGGCGGGGAGCTGGAGAGAGCTCCTGTGATTTCTACCCAGGCGCTTGGTGGCTTCTCT	1620
Db	1562	GCGGCGGGGAGCTGGAGAGAGCTCCTGTGATTTCTACCCAGGCGCTTGGTGGCTTCTCT	1621
Qy	1621	GATGTGATCTCATTTCTGAGAGCTTCTGAAGCTTGGGGGGGGGGGGGGGGGGGGGGGGGGGG	1680
Db	1622	GATGTGATCTCATTTCTGAGAGCTTCTGAAGCTTGGGGGGGGGGGGGGGGGGGGGGGGGGGG	1681
Qy	1681	GGCTTCCCTCAGTGACCTCTCTCTGTGTCAGCAGCGGGGGGGGGGGGGGGGGGGGGGGGGGG	1740

Db 1682 GGCCTTCCCTCAGTACCCCTCATCTCCCTGTCAGCAGCCAGGGGCCCCCAGGCTGGAGGCG 1741
Qy 1741 AGCCATTGTAGAGCCAGAGGGGAACCACTTTTGGGAACCCCAACCCCTCCATGTGATGGA 1800
Db 1742 AGCCATTGTAGAGCCAGAGGGGAACCACTTTTGGGAACCCCAACCCCTCCATGTGATGGA 1801
Qy 1801 GAACTGTCTGAGGGCAGAGGGATCTACGCCAGCAGGTGGAGGCTTGTTCAGGGGGTGGC 1860
Db 1802 GAACTGTCTGAGGGCAGAGGGATCTACGCCAGCAGGTGGAGGCTTGTTCAGGGGGTGGC 1861
Qy 1861 GGCCTTTACGCTCTGGCTTGGCCTTTGCTTCACACGTG 1899
Db 1862 GGCCTTTACGCTCTGGCTTGGCCTTTGCTTCACACGTG 1900

RESULT 4
US-10-664-667-4
; Sequence 4, Application US/10664667
; Publication No. US20040058377A1
; GENERAL INFORMATION:
; APPLICANT: Gu, Wei
; TITLE OF INVENTION: NOVEL G-PROTEIN COUPLED RECEPTORS AND USES THEREFOR
; CURRENT APPLICATION NUMBER: US/10/664,667
; CURRENT FILING DATE: 2003-09-18
; PRIOR APPLICATION NUMBER: US/09/851,595
; PRIOR FILING DATE: 2000-05-08
; PRIOR APPLICATION NUMBER: 09/556,588
; PRIOR FILING DATE: 2000-05-08
; PRIOR APPLICATION NUMBER: 60/132,896
; PRIOR FILING DATE: 1999-05-06
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 4
; LENGTH: 2486
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (2)..(1900)
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (172)
; OTHER INFORMATION: n = any nucleotide
US-10-664-667-4

Query Match 99.9%; Score 1898; DB 17; Length 2486;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1899; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 AATAGACTCACTATAGGGAAGCTGGTACGCTGACGCTACCGGTCCGGGAATTCCTCCGGG 60
Db 2 AATAGACTCACTATAGGGAAGCTGGTACGCTGACGCTACCGGTCCGGGAATTCCTCCGGG 61
Qy 61 TCGACCCACGGGTCGGTGGAGGGAGCCAGAGGTCTGAGCCTGCCGGCTCATCCAGCCCTCT 120
Db 62 TCGACCCACGGGTCGGTGGAGGGAGCCAGAGGTCTGAGCCTGCCGGCTCATCCAGCCCTCT 121
Qy 121 CTTGCTGCCCTAGCGGCTCCAAACAAACCGCATCTGGGAAATTGGAGCTNGACACCTTC 180
Db 122 CTTGCTGCCCTAGCGGCTCCAAACAAACCGCATCTGGGAAATTGGAGCTNGACACCTTC 181
Qy 181 AGCCAGCTGAGCTCCCTCGAAGCCCTGGATCTTAGCTGGAACGCCATCCGGTCCATCCAC 240
Db 182 AGCCAGCTGAGCTCCCTCGAAGCCCTGGATCTTAGCTGGAACGCCATCCGGTCCATCCAC 241
Qy 241 CTTGAGGCTCTTCCACCCCTGACCTCCCTGTGCTCAAGCTGGACCTGACAGACAAACAGCTG 300
Db 242 CTTGAGGCTCTTCCACCCCTGACCTCCCTGTGCTCAAGCTGGACCTGACAGACAAACAGCTG 301
Qy 301 ACCACACTGCCCCCTGGCTGGAGCTTGGGGCTTGATGCACTCTGAAGCTCAAAAGGGAACCTT 360

Db 302 ACCACACTGCCCCCTGGCTGGAGCTTGGGGGCTTGATGCACTCTGAAGCTCAAAAGGGAACCTT 361
Qy 361 GCTCTCTCCAGGGCTTCTCCAAAGGACAGTATTTCCAAAGCTGAGGATCTCTGAGGTGCT 420
Db 362 GCTCTCTCCAGGGCTTCTCCAAAGGACAGTATTTCCAAAGCTGAGGATCTCTGAGGTGCT 421
Qy 421 TATGCTTACCAAGTGTCTCTTATGGGATGTGTCAGCTTCTTCAAGGCTCTCTGGGAG 480
Db 422 TATGCTTACCAAGTGTCTCTTATGGGATGTGTCAGCTTCTTCAAGGCTCTCTGGGAG 481
Qy 481 TGGGAGGCTGAAGACCTTCACTTGTATGATGAGGAGTCTTCAAAAGGCTCTCTGGGCTC 540
Db 482 TGGGAGGCTGAAGACCTTCACTTGTATGATGAGGAGTCTTCAAAAGGCTCTCTGGGCTC 541
Qy 541 CTTGCCAGACAAAGCAGAGAACCACTATGACAGAGACCTGGATGAGCTCCAGCTGGAGATG 600
Db 542 CTTGCCAGACAAAGCAGAGAACCACTATGACAGAGACCTGGATGAGCTCCAGCTGGAGATG 601
Qy 601 GAGGACTCAAAAGCCACACCCAGTGTCCAGTGTAGCCCTTACTTCCAGGCCCCCTTCAAGGCC 660
Db 602 GAGGACTCAAAAGCCACACCCAGTGTCCAGTGTAGCCCTTACTTCCAGGCCCCCTTCAAGGCC 661
Qy 661 TGTGAGTACCTCTTTGAAAGCTGGGCACTCCGCTGGCCGCTGGCCGCTATCGTGTGCTC 720
Db 662 TGTGAGTACCTCTTTGAAAGCTGGGCACTCCGCTGGCCGCTGGCCGCTATCGTGTGCTC 721
Qy 721 TCCGTGCTCTGCAATGGACTGCTGCTGACCCGCTTTCGCTGGCGGCTTCCGCCCCCTG 780
Db 722 TCCGTGCTCTGCAATGGACTGCTGCTGACCCGCTTTCGCTGGCGGCTTCCGCCCCCTG 781
Qy 781 CCCCCTGCTCAAGTTGTGGTAGTGCAGTGCAGGCGCCAAACACTTCACTGGGCAATTTCC 840
Db 782 CCCCCTGCTCAAGTTGTGGTAGTGCAGTGCAGGCGCCAAACACTTCACTGGGCAATTTCC 841
Qy 841 TGTGGCTCTTAGCTCAGTGCATGCCCTGACCTTTGGTCACTTCTCTGAGTACGAGGCC 900
Db 842 TGTGGCTCTTAGCTCAGTGCATGCCCTGACCTTTGGTCACTTCTCTGAGTACGAGGCC 901
Qy 901 CGCTGGGAGACGGGCTAGGCTGCGGGCCACTGGCTTCCCTGGGAGTACTTTGGGCTGGAG 960
Db 902 CGCTGGGAGACGGGCTAGGCTGCGGGCCACTGGCTTCCCTGGCAGTACTTTGGGCTGGAG 961
Qy 961 GCATCGGTGCTGCTCACTCTTGGCGCAGTGCAGTGCAGCGTCTCCGTCTCTGTGCTC 1020
Db 962 GCATCGGTGCTGCTCACTCTTGGCGCAGTGCAGTGCAGCGTCTCCGTCTCTGTGCTC 1021
Qy 1021 CGGCTCATGGGAAGTCCCCCTCCCTGGGAGCGTTTCGAGCAGGGGCTCTTAGGCTGGCTG 1080
Db 1022 CGGCTCATGGGAAGTCCCCCTCCCTGGGAGCGTTTCGAGCAGGGGCTCTTAGGCTGGCTG 1081
Qy 1081 GCATGGCAGGGCTGGCGCGCCACTGCCCCCTGGCCTCAGTGGGAGAAATACGGGCTTCC 1140
Db 1082 GCATGGCAGGGCTGGCGCGCCACTGCCCCCTGGCCTCAGTGGGAGAAATACGGGCTTCC 1141
Qy 1141 CCATCTGCTGCTCCTACGCGCCACTTGAAGGCTCAGCCAGCAGCCCTGGGCTTCAACCGTG 1200
Db 1142 CCATCTGCTGCTCCTACGCGCCACTTGAAGGCTCAGCCAGCAGCCCTGGGCTTCAACCGTG 1201
Qy 1201 GCTTGGTGTATGAACTCCTTCTGTCTTCTGTGCTGGCGGCTGCTATCATCAAACTG 1260
Db 1202 GCTTGGTGTATGAACTCCTTCTGTCTTCTGTGCTGGCGGCTGCTATCATCAAACTG 1261
Qy 1261 TACTGTGACCTGCGCGGGGAGCTTTGAGGCGCTGTGGGACTCGCCCATGGTGGGAGC 1320
Db 1262 TACTGTGACCTGCGCGGGGAGCTTTGAGGCGCTGTGGGAGCTGGGAGCTGGTGGGAGC 1321
Qy 1321 GTGCTTGGCTCATCTTTCGACAGCGGCTCCTTACTGTCTCCGTGGGCTTCTCTCAGCTTC 1380
Db 1322 GTGCTTGGCTCATCTTTCGACAGCGGCTCCTTACTGTCTCCGTGGGCTTCTCTCAGCTTC 1381
Qy 1381 GCCTTCATGTGGGCTCTTCCCTGTGTCAGGCCCGGCTCAAGTCTGTCTCTGTGTG 1440
Db 1382 GCCTTCATGTGGGCTCTTCCCTGTGTCAGGCCCGGCTCAAGTCTGTCTCTGTGTG 1441

QY 1441 GTGCTGCCCTGCTGCTGCTCAACCCACTGCTGTACCTGCTCTTCAACCCCACTTC 1500
 Db 1442 GTGCTGCCCTGCTGCTGCTCAACCCACTGCTGTACCTGCTCTTCAACCCCACTTC 1501
 QY 1501 CGGGATGACCTTTCGGCGGCTTCGGCGGCGCGAGGGAGTCAAGGCGGCTTACGCTATGCT 1560
 Db 1502 CGGGATGACCTTTCGGCGGCTTCGGCGGCGCGAGGGAGTCAAGGCGGCTTACGCTATGCT 1561
 QY 1561 GCGGCGGGAGCTGGAGAGAGCTCTGTGATTTACCAAGGCGGCTTACGCTATGCT 1620
 Db 1562 GCGGCGGGAGCTGGAGAGAGCTCTGTGATTTACCAAGGCGGCTTACGCTATGCT 1621
 QY 1621 GATGTGATCTCATTTCTGAAAGCTTCTGAAGCTGGCGGCGGCTTACGCTATGCT 1680
 Db 1622 GATGTGATCTCATTTCTGAAAGCTTCTGAAGCTGGCGGCGGCTTACGCTATGCT 1681
 QY 1681 GGCCTTCCCTCAGTGACCTCTCATCTCTGTGAGAGCGGCGGCGGCTTACGCTATGCT 1740
 Db 1682 GGCCTTCCCTCAGTGACCTCTCATCTCTGTGAGAGCGGCGGCGGCTTACGCTATGCT 1741
 QY 1741 AGCCATTGTGAGAGCGAGAGGGAACCACTTTGGGAAACCCCAACCTTCAATGATGGA 1800
 Db 1742 AGCCATTGTGAGAGCGAGAGGGAACCACTTTGGGAAACCCCAACCTTCAATGATGGA 1801
 QY 1801 GAACTGCTGTGAGGCGAGAGGATCTACGCGAGCGAGGCTTGTCAAGGGGTGGC 1860
 Db 1802 GAACTGCTGTGAGGCGAGAGGATCTACGCGAGCGAGGCTTGTCAAGGGGTGGC 1861
 QY 1861 GGCCTTACGCGCTCTGGCTTGGCTTTGCTTTCACACGCTG 1899
 Db 1862 GGCCTTACGCGCTCTGGCTTGGCTTTGCTTTCACACGCTG 1900

RESULT 5

US-09-851-595-9
 ; Sequence 9, Application US/09851595
 ; Publication No. US20030166047A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Gu, Wei
 ; TITLE OF INVENTION: NOVEL G-PROTEIN COUPLED RECEPTORS AND USES THEREFOR
 ; FILE REFERENCE: MNI-080CP
 ; CURRENT APPLICATION NUMBER: US/09/851,595
 ; PRIOR FILING DATE: 2000-05-08
 ; PRIOR APPLICATION NUMBER: 09/556,588
 ; PRIOR FILING DATE: 2000-05-08
 ; PRIOR APPLICATION NUMBER: 60/132,896
 ; PRIOR FILING DATE: 1999-05-06
 ; NUMBER OF SEQ ID NOS: 12
 ; SOFTWARE: PatentIn Ver. 2.0
 ; SEQ ID NO 9
 ; LENGTH: 2208
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; FEATURE:
 ; NAME/KEY: CDS
 ; LOCATION: (1)...(2208)
 US-09-851-595-9

Query Match 92.4%; Score 1754.4; DB 10; Length 2208;
 Best Local Similarity 99.9%; Pred. No. 0;
 Matches 1766; Conservative 0; Mismatches 1; Indels 1; Gaps 1;
 QY 132 AGCGGCTCCAAACAACCGCATCTGGAAATTGGAGCTNGACACCTTACGCCAGCTGAG 191
 Db 442 ATCGGCTCCAAACAACCGCATCTGGAAATTGGAGCT-GACACCTTACGCCAGCTGAG 500
 QY 192 CTCCTCGAAGCCCTGATCTTACCTGGAAGCCATCGGTCCATCCACCTGAGGCTT 251
 Db 501 CTCCTCGAAGCCCTGATCTTACCTGGAAGCCATCGGTCCATCCACCTGAGGCTT 560
 QY 252 CTCACCTGACCTCCCTGCTCAAGCTGGAGCTGACAGAAACACGAGTGCACCTGCC 311

Db 561 CTCCACCCCTGCACCTCCCTGGTCAAGCTGGAGCTTGACAGAAACACGAGTGACCACTGCC 620
 QY 312 CCTGGCTGGACTTGGGGGCTTGATGATCTGAAGCTCAAAGGGAACCTTCTCTCTCCCA 371
 Db 621 CCTGGCTGGACTTGGGGGCTTGATGATCTGAAGCTCAAAGGGAACCTTCTCTCTCCCA 680
 QY 372 GGCCTTCTCCAAAGGACAGTTTCCCAAACTGAGGATCTTGAGGTGCTTATGCTTACCA 431
 Db 681 GGCCTTCTCCAAAGGACAGTTTCCCAAACTGAGGATCTTGAGGTGCTTATGCTTACCA 740
 QY 432 GTGCTGTCCTATGAGGATGTGTGACAGCTTCTTCAAGGCGCTCTGGGAGTGGGAGGTGA 491
 Db 741 GTGCTGTCCTATGAGGATGTGTGACAGCTTCTTCAAGGCGCTCTGGGAGTGGGAGGTGA 800
 QY 492 AGACCTTCACTTGAATGATGAGGATCTTCAAAGAGCCCTTGGGCGCTCTTGGCAGACA 551
 Db 801 AGACCTTCACTTGAATGATGAGGATCTTCAAAGAGCCCTTGGGCGCTCTTGGCAGACA 860
 QY 552 AGCAGAGAAACACTATGACAGGACCTGGATGAGTCCAGCTGGAGATGGAGGACTCAAA 611
 Db 861 AGCAGAGAAACACTATGACAGGACCTGGATGAGTCCAGCTGGAGATGGAGGACTCAAA 920
 QY 612 GCCACACCCAGTGTCCAGTGTAGCCCTACTCCAGGCGCCTTCAAGCGCTGTGAGTACCT 671
 Db 921 GCCACACCCAGTGTCCAGTGTAGCCCTACTCCAGGCGCCTTCAAGCGCTGTGAGTACCT 980
 QY 672 CTTTGAAGCTGGGGCATCCGCTGGCGGTGGGGCCATCGTGTGCTCTCCGTGCTGTG 731
 Db 981 CTTTGAAGCTGGGGCATCCGCTGGCGGTGGGGCCATCGTGTGCTCTCCGTGCTGTG 1040
 QY 732 CAATGGAATCGTGTGCTGACCGTGTTCGCTGGCGGCGCTGCCCGCTGCCCGGTCAA 791
 Db 1041 CAATGGAATCGTGTGCTGACCGTGTTCGCTGGCGGCGCTGCCCGCTGCCCGGTCAA 1100
 QY 792 GTTTGTGAGTGCGAATTGACAGCGCCAAACACTTGACTGGCAATTTCTGTGCGCTTCT 851
 Db 1101 GTTTGTGAGTGCGAATTGACAGCGCCAAACACTTGACTGGCAATTTCTGTGCGCTTCT 1160
 QY 852 AGCCTCAGTGCATGCCCTGACCTTTGCTCAGTTCTCTGAGTACGAGGCGGCTGGGAGAC 911
 Db 1161 AGCCTCAGTGCATGCCCTGACCTTTGCTCAGTTCTCTGAGTACGAGGCGGCTGGGAGAC 1220
 QY 912 GGGCTAGGCTGCGGGCCACTGGCTTCTGCGAGTACTTTGGGTGGAGGACATCGGTGCT 971
 Db 1221 GGGCTAGGCTGCGGGCCACTGGCTTCTGCGAGTACTTTGGGTGGAGGACATCGGTGCT 1280
 QY 972 GCTGCTCACTCTGCGCGCAGTGCAGTGCAGCGTCTCCGTCTCCTGTGTCGCGGCTATGG 1031
 Db 1281 GCTGCTCACTCTGCGCGCAGTGCAGTGCAGCGTCTCCGTCTCCTGTGTCGCGGCTATGG 1340
 QY 1032 GAAGTCCCTCTCCCTGGGAGCGTTTCGAGCAGGCGGCTTACGCTGCTGGCAGTGGGAGG 1091
 Db 1341 GAAGTCCCTCTCCCTGGGAGCGTTTCGAGCAGGCGGCTTACGCTGCTGGCAGTGGGAGG 1400
 QY 1092 GCTGGCGCGCAGTGCCTGCGCTCAGTGGAGAAATACGGGCGCTTCCCACTCTCCCT 1151
 Db 1401 GCTGGCGCGCAGTGCCTGCGCTCAGTGGAGAAATACGGGCGCTTCCCACTCTCCCT 1460
 QY 1152 GCCTTAGCGGCCACCTGAGGGTCAAGCAGCAGCGCTGGGCTTCAACCGTGGCGCTGTGAT 1211
 Db 1461 GCCTTAGCGGCCACCTGAGGGTCAAGCAGCAGCGCTGGGCTTCAACCGTGGCGCTGTGAT 1520
 QY 1212 GATGAATCTCTCTGTTTCTGTCGTCGCGGTGCTCATCAAACTGACTGTGACCT 1271
 Db 1521 GATGAATCTCTCTGTTTCTGTCGTCGCGGTGCTCATCAAACTGACTGTGACCT 1580
 QY 1272 GCCGCGGGGCGACTTTTGGAGCGGCTGGGACTCGCGCATGCTGAGGACGCTGGCGCTGGCT 1331
 Db 1581 GCCGCGGGGCGACTTTTGGAGCGGCTGGGACTCGCGCATGCTGAGGACGCTGGCGCTGGCT 1640
 QY 1332 CATCTTTCGACAGCGGCTCCTTACTGTCCCGTGGCGCTTCTCAGCTTCCGCTCCATGCT 1391
 Db 1641 CATCTTTCGACAGCGGCTCCTTACTGTCCCGTGGCGCTTCTCAGCTTCCGCTCCATGCT 1700

Db 1521 GATGAACCTCTCTCTGTTCTCTGCTGCTGCGCGGTGCTTACATCAAACTGTACTGTGACCT 1580
Qy 1272 GCGCGGGGCGACTTTGAGGCGCGTGTGGGACTGCGGCCATGGTGAGGCACGCTGGCCTGGCT 1331
Db 1581 GCGCGGGGCGACTTTGAGGCGCGTGTGGGACTGCGGCCATGGTGAGGCACGCTGGCCTGGCT 1640
Qy 1332 CATCTTCGAGAGCGGGCTCTCTACTGTCCCGTGGCCTTCCTCAAGCTTCGCTTCATGCT 1391
Db 1641 CATCTTCGAGAGCGGGCTCTCTACTGTCCCGTGGCCTTCCTCAAGCTTCGCTTCATGCT 1700
Qy 1392 GGGCCTCTTCCCTGTCAAGCGCGGCGGTCAAGTCTGCTGCTGCTGCTGCTGCTGCTGCT 1451
Db 1701 GGGCCTCTTCCCTGTCAAGCGCGGCGGTCAAGTCTGCTGCTGCTGCTGCTGCTGCTGCT 1760
Qy 1452 GCCTGCTGCTCAACCCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1511
Db 1761 GCCTGCTGCTCAACCCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1820
Qy 1512 TCGGCGGCTTCGGCCCGCGAGGGAATCAAGGCGGCTGAGGCTGCTGCTGCTGCTGCTGCTGCT 1571
Db 1821 TCGGCGGCTTCGGCCCGCGAGGGAATCAAGGCGGCTGAGGCTGCTGCTGCTGCTGCTGCTGCT 1880
Qy 1572 GCTGAGAGAGCTCTCTGATTTCTACCCAGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1631
Db 1881 GCTGAGAGAGCTCTCTGATTTCTACCCAGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1940
Qy 1632 CATCTGGAAGCTTCTGAAGCTGGGCGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1691
Db 1941 CATCTGGAAGCTTCTGAAGCTGGGCGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2000
Qy 1692 AGTGACCTCATCTCTGCTGAGCAGCAGGCGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1751
Db 2001 AGTGACCTCATCTCTGCTGAGCAGCAGGCGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2060
Qy 1752 AGAGCCAGAGGGGACCACTTTGGGAACCCCAACCCCTCCATGGATGAGAACTGCTGCT 1811
Db 2061 AGAGCCAGAGGGGACCACTTTGGGAACCCCAACCCCTCCATGGATGAGAACTGCTGCT 2120
Qy 1812 GAGGCGAGAGGATCTAGCCAGCAGGTGGAGGCTTGTGAGGGGTGGCGGCTTTTCAGCC 1871
Db 2121 GAGGCGAGAGGATCTAGCCAGCAGGTGGAGGCTTGTGAGGGGTGGCGGCTTTTCAGCC 2180
Qy 1872 CTCTGGCTTGGCTTTGCTTTCACAGTG 1899
Db 2181 CTCTGGCTTGGCTTTGCTTTCACAGTG 2208

RESULT 8

US-10-664-667-7
; Sequence 7, Application US/10664667
; Publication No. US20040058377A1
; GENERAL INFORMATION:
; APPLICANT: Gu, Wei
; TITLE OF INVENTION: NOVEL G-PROTEIN COUPLED RECEPTORS AND USES THEREFOR
; FILE REFERENCE: MNI-080CP
; CURRENT APPLICATION NUMBER: US/10/664,667
; CURRENT FILING DATE: 2003-09-18
; PRIOR APPLICATION NUMBER: US/09/851,595
; PRIOR FILING DATE: 2000-05-08
; PRIOR APPLICATION NUMBER: 09/556,588
; PRIOR FILING DATE: 2000-05-08
; PRIOR APPLICATION NUMBER: 60/132,896
; PRIOR FILING DATE: 1999-05-06
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: PatentIn ver. 2.0
; SEQ ID NO 7
; LENGTH: 2711
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(2208)

US-10-664-667-7

Query Match 92.4%; Score 1754.4; DB 17; Length 2711;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1766; Conservative 0; Mismatches 1; Indels 1; Gaps 1;
Qy 132 AGCGGCTTCCAAACACACCGCATCTGGGAAATTTGAGGTNGACACTTTCAGCCAGCTGAG 191
Db 442 ATCGGCTTCCAAACACACCGCATCTGGGAAATTTGAGGT-GACACCTTTCAGCCAGCTGAG 500
Qy 192 CTCCTGCAAGCCCTGGATCTTAGCTGGAACGCCATCCGGTCCATCCACCTTCAGGCGCTT 251
Db 501 CTCCTGCAAGCCCTGGATCTTAGCTGGAACGCCATCCGGTCCATCCACCTTCAGGCGCTT 560
Qy 252 CTCACCCCTGCATCCCTGGTCAAGCTGGAACCTGACAGAACACCAAGCTGACACACTGCG 311
Db 561 CTCACCCCTGCATCCCTGGTCAAGCTGGAACCTGACAGAACACCAAGCTGACACACTGCG 620
Qy 312 CTTGGCTGGAATTTGGGGCTTGTGATCTGAAAGCTCAAGGGAACCTTGTCTCTCCCA 371
Db 621 CTTGGCTGGAATTTGGGGCTTGTGATCTGAAAGCTCAAGGGAACCTTGTCTCTCCCA 680
Qy 372 GGCCTTCTCAAGGACAGCTTCCCAAACTCAGGATCTCGAGGTGCTTATGCGCTACCA 431
Db 681 GGCCTTCTCAAGGACAGCTTCCCAAACTCAGGATCTCGAGGTGCTTATGCGCTACCA 740
Qy 432 GTGCTGTCCCTATGGGATGTGTGCGAGCTTCTTCAAGGCTCTTGGGCACTGGAGGCTGA 491
Db 741 GTGCTGTCCCTATGGGATGTGTGCGAGCTTCTTCAAGGCTCTTGGGCACTGGAGGCTGA 800
Qy 492 AGACCTTCACTTGTATGATGAGGAGTCTTCAAAAGGCGCTTGGGCTCTTGTCCAGACA 551
Db 801 AGACCTTCACTTGTATGATGAGGAGTCTTCAAAAGGCGCTTGGGCTCTTGTCCAGACA 860
Qy 552 AGCAGAGAACCACTATGACACAGGACCTGGATGAGCTCCAGCTGGAGATGGAGGACTCAA 611
Db 861 AGCAGAGAACCACTATGACACAGGACCTGGATGAGCTCCAGCTGGAGATGGAGGACTCAA 920
Qy 612 GCCACACCCAGTGTCCAGTGTAGCCCTTCTTCAAGGCGCTTCAAGCCCTGTGAGTACT 671
Db 921 GCCACACCCAGTGTCCAGTGTAGCCCTTCTTCAAGGCGCTTCAAGCCCTGTGAGTACT 980
Qy 672 CTTTGAAGCTGGGGCATCCGCTTGGCGCTGTGGGCCATCGTGTGCTCTCGTGTCTG 731
Db 981 CTTTGAAGCTGGGGCATCCGCTTGGCGCTGTGGGCCATCGTGTGCTCTCGTGTCTG 1040
Qy 732 CAATGGACTGTGCTGACCGTGTTCGCTGGGGGCTGCGCCCTTGGCCCGGCTCAA 791
Db 1041 CAATGGACTGTGCTGACCGTGTTCGCTGGGGGCTGCGCCCTTGGCCCGGCTCAA 1100
Qy 792 GTTTGTGTAGTGTGCGATTGCGAGCGCCCAACACCTTGAAGCTGGCATTTCTGTGGCTTCT 851
Db 1101 GTTTGTGTAGTGTGCGATTGCGAGCGCCCAACACCTTGAAGCTGGCATTTCTGTGGCTTCT 1160
Qy 852 AGCTCAGTCGATCCCTGACCTTTTGTGTCAGTCTCTGAGTACGAGCCCTGCGAGAC 911
Db 1161 AGCTCAGTCGATCCCTGACCTTTTGTGTCAGTCTCTGAGTACGAGCCCTGCGAGAC 1220
Qy 912 GGGCTAGGCTGCGGGGCACTGCTTCTGCGAGTACTTGGGTGCGAGGATCGGTGCT 971
Db 1221 GGGCTAGGCTGCGGGGCACTGCTTCTGCGAGTACTTGGGTGCGAGGATCGGTGCT 1280
Qy 972 GCTGCTACTCTGCGCGCAGTGCAGGCTCTCGGTCTCTGCTGCTGCTGCTGCTGCTGCT 1031
Db 1281 GCTGCTACTCTGCGCGCAGTGCAGGCTCTCGGTCTCTGCTGCTGCTGCTGCTGCTGCT 1340
Qy 1032 GAAGTCCCTCTCCTTGGGCGAGCTTCGAGCAGGGGTCTTAGGCTGCTTGGCACTGGCAGG 1091
Db 1341 GAAGTCCCTCTCCTTGGGCGAGCTTCGAGCAGGGGTCTTAGGCTGCTTGGCACTGGCAGG 1400
Qy 1092 GCTGGCGCGCACTGCGCCCTTGGCTCAGTGGGAGAAATACGGGGCTTCCCACTCTGCT 1151
Db 1401 GCTGGCGCGCACTGCGCCCTTGGCTCAGTGGGAGAAATACGGGGCTTCCCACTCTGCT 1460

Qy 1152 GCCCTACGCGCACTGAGGCTCAGCCAGCAGCCCTGGGCTTCCCGTGGCCCTGGTGAT 1211
Db 1461 GCCCTACGCGCACTGAGGCTCAGCCAGCAGCCCTGGGCTTCCCGTGGCCCTGGTGAT 1520
Qy 1212 GATGAATCCTCTCTGTTTCTGGTCTGGCGGTGCTTACATCAAACTGTACTGTGACCT 1271
Db 1521 GATGAATCCTCTCTGTTTCTGGTCTGGCGGTGCTTACATCAAACTGTACTGTGACCT 1580
Qy 1272 GCCGCGGCGCACTTTGAGGCGGTGTGGGACTGGCCATGTGTAGGCACTGTGGCTGGCT 1331
Db 1581 GCCGCGGCGCACTTTGAGGCGGTGTGGGACTGGCCATGTGTAGGCACTGTGGCTGGCT 1640
Qy 1332 CATCTTCGACAGCGGCTCCTACTGTCGCGTGGCTTCTCAGCTTCGCTCCATGCT 1391
Db 1641 CATCTTCGACAGCGGCTCCTACTGTCGCGTGGCTTCTCAGCTTCGCTCCATGCT 1700
Qy 1392 GGGCTCTTCCCTGTACGCCCGAGGCGCTCAAGTCTGTCTCTGCTGGTGGTGGCTCCCT 1451
Db 1701 GGGCTCTTCCCTGTACGCCCGAGGCGCTCAAGTCTGTCTCTGCTGGTGGTGGCTCCCT 1760
Qy 1452 GCCTGCTGCTCAACCCACTGCTGTACTGCTCTTTCAACCCCACTTCCGGGATGACCT 1511
Db 1761 GCCTGCTGCTCAACCCACTGCTGTACTGCTCTTTCAACCCCACTTCCGGGATGACCT 1820
Qy 1512 TCGCGGGCTTCGGCCCGCGCAGGGGACTCAGGGCCCTAGCCCTATGCTGCGCGCGGGA 1571
Db 1821 TCGCGGGCTTCGGCCCGCGCAGGGGACTCAGGGCCCTAGCCCTATGCTGCGCGCGGGA 1880
Qy 1572 GCTGAGAGAGCTCCTGTGATTTACCCAGGCGCTTGTAGCCCTTCTCTGATGTGGATCT 1631
Db 1881 GCTGAGAGAGCTCCTGTGATTTACCCAGGCGCTTGTAGCCCTTCTCTGATGTGGATCT 1940
Qy 1632 CATCTGGAAGCTTCTGAAGCTGGCGCGCCCTTGGGCTGGAGACCTATGCTTCCCTC 1691
Db 1941 CATCTGGAAGCTTCTGAAGCTGGCGCGCCCTTGGGCTGGAGACCTATGCTTCCCTC 2000
Qy 1692 AGTGACCTCATCTCTGTGAGCAGCAGGGGCCCCCAGGCTGGAGGCGACCACTTGTGT 1751
Db 2001 AGTGACCTCATCTCTGTGAGCAGCAGGGGCCCCCAGGCTGGAGGCGACCACTTGTGT 2060
Qy 1752 AGAGCAGAGGGGAACCACTTTTGGGAACCCCAACCCCTCCATGGATGGAGAACTGCTGT 1811
Db 2061 AGAGCAGAGGGGAACCACTTTTGGGAACCCCAACCCCTCCATGGATGGAGAACTGCTGT 2120
Qy 1812 GAGGCGAGAGGGATCTAGCCAGCAGGTGGAGGCTTGTTCAGGGGGTGGCGGCTTTCAGCC 1871
Db 2121 GAGGCGAGAGGGATCTAGCCAGCAGGTGGAGGCTTGTTCAGGGGGTGGCGGCTTTCAGCC 2180
Qy 1872 CTCTGGCTTGGCTTTCCTTTCACACGTG 1899
Db 2181 CTCTGGCTTGGCTTTCCTTTCACACGTG 2208

RESULT 9
US-09-851-595-12
; Sequence 12, Application US/09851595
; Publication No. US20030166047A1
; GENERAL INFORMATION:
; APPLICANT: Gu, Wei
; TITLE OF INVENTION: NOVEL G-PROTEIN COUPLED RECEPTORS AND USES THEREFOR
; FILE REFERENCE: MNI-080CP
; CURRENT APPLICATION NUMBER: US/09/851,595
; CURRENT FILING DATE: 2000-05-08
; PRIOR APPLICATION NUMBER: 09/556,588
; PRIOR FILING DATE: 2000-05-08
; PRIOR APPLICATION NUMBER: 60/132,896
; PRIOR FILING DATE: 1999-05-06
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: PatentIn ver. 2.0
; SEQ ID NO 12
; LENGTH: 2901
; TYPE: DNA

; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)... (2901)
US-09-851-595-12

Query Match 92.4%; Score 1754.4; DB 10; Length 2901;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1766; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

Qy 132 AGCGGCTCCAAACAACACCGCATCTGGGAAATTGGAGCTNGACACCTTCAGCCAGCTGAG 191
Db 1135 ATCGGCTCCAAACAACACCGCATCTGGGAAATTGGAGCT-GACACCTTCAGCCAGCTGAG 1193
Qy 192 CTCCCTGCAAGCCCTGATCTTAGCTGGAAGGCCATCGGCTCCATCCACCTGAGGCTT 251
Db 1194 CTCCCTGCAAGCCCTGATCTTAGCTGGAAGGCCATCCGGTCCATCCACCTGAGGCTT 1253
Qy 252 CTCCACCTTCACCTCCCTGGTCAAGCTGGAGCTTGACAGACAACCCAGCTGACCACTGCC 311
Db 1254 CTCCACCTTCGACTCCCTCCCTGGTCAAGCTGGAGCTTGACAGACAACCCAGCTGACCACTGCC 1313
Qy 312 CTGCTGCTGACTTTGGGGCTTGTATGATCTGAAGCTCAAAAGGAACTTGTCTCTCCCA 371
Db 1314 CTGCTGCTGACTTTGGGGCTTGTATGATCTGAAGCTCAAAAGGAACTTGTCTCTCCCA 1373
Qy 372 GGCCTTCTCAAGAGCAGTTTCCCAAACTGAGGATCTTGGAGGTGCTTATGCTTACCA 431
Db 1374 GGCCTTCTCAAGAGCAGTTTCCCAAACTGAGGATCTTGGAGGTGCTTATGCTTACCA 1433
Qy 432 GTGCTGCTCCCTATGGGATGTGCCAGCTTCTTCAAGGCCCTCTGGGAGTGGGAGGCTGA 491
Db 1434 GTGCTGCTCCCTATGGGATGTGTGCCAGCTTCTTCAAGGCCCTCTGGGAGTGGGAGGCTGA 1493
Qy 492 AGACCTTACCTTGTATGATGAGGAGTCTTCAAAAGGCCCTCTGGGCTCTCTTGCAGACA 551
Db 1494 AGACCTTACCTTGTATGATGAGGAGTCTTCAAAAGGCCCTCTGGGCTCTCTTGCAGACA 1553
Qy 552 AGCAGAGAACCACTATGACAGGACCTGGATGAGCTCCAGCTGGAGATGGAGNCTCAA 611
Db 1554 AGCAGAGAACCACTATGACAGGACCTGGATGAGCTCCAGCTGGAGATGGAGNCTCAA 1613
Qy 612 GCACACCCAGTGTCCAGTGTAGCCCTACTCCAGGCCCTTCAAGGCCCTGTGAGTACCT 671
Db 1614 GCACACCCAGTGTCCAGTGTAGCCCTACTCCAGGCCCTTCAAGGCCCTGTGAGTACCT 1673
Qy 672 CTTTGAAGCTGGGGCATCCGCTGGCCGTGTGGGCCATCCTGTGTCTCCGTGCTCTG 731
Db 1674 CTTTGAAGCTGGGGCATCCGCTGGCCGTGTGGGCCATCCTGTGTCTCCGTGCTCTG 1733
Qy 732 CAATGGACTGTGCTGCTGACCGTGTGGTGGGGGCTTGGCCCTTGGCCCTGGCCCTGCA 791
Db 1734 CAATGGACTGTGCTGCTGACCGTGTGGTGGGGGCTTGGCCCTTGGCCCTGGCCCTGCA 1793
Qy 792 GTTTGTGTAGGTGCGATTTCAGCGCCAAACACTTCTGACTGGCATTTCTGTGGCTTCT 851
Db 1794 GTTTGTGTAGGTGCGATTTCAGCGCCAAACACTTCTGACTGGCATTTCTGTGGCTTCT 1853
Qy 852 AGCCTCAGTGCATGCCCTGACCTTTTGGTCAAGTTCTCTGAGTACGGAGCCCCCTGGAGAC 911
Db 1854 AGCCTCAGTGCATGCCCTGACCTTTTGGTCAAGTTCTCTGAGTACGGAGCCCCCTGGAGAC 1913
Qy 912 GGGCTAGGCTGCGGGCCACTGGCTTCTGGCAGTACTTTGGGTGCGAGGATCGGTGCT 971
Db 1914 GGGCTAGGCTGCGGGCCACTGGCTTCTGGCAGTACTTTGGGTGCGAGGATCGGTGCT 1973
Qy 972 GCTGCTCACTCTGCGCGCAGTGCAGTGCAGCTCTCCGTCTCCTGTCTCCGGGCTATGG 1031
Db 1974 GCTGCTCACTCTGCGCGCAGTGCAGTGCAGCTCTCCGTCTCCTGTCTCCGGGCTATGG 2033
Qy 1032 GAAATCCCCCTCCCTGGGAGCGTTCGAGCAGGGGCTCTAGGCTGCTGACCTGGCAGG 1091
Db 2034 GAAATCCCCCTCCCTGGGAGCGTTCGAGCAGGGGCTCTAGGCTGCTGACCTGGCAGG 2093

1542	GCCACACCCCAAGTGTCCAGTGTAGCCCTACTCCAGGGCCCCCTTCAAGGCCCTGTGTAGTACCT	1601
Qy	CTTTGAAAGCTGGGGCAATCCGCCCTGGCGGTGTGGGCCCATCGTGTGTCTTCCCGTCTCTG	731
Db	CTTTGAAAGCTGGGGCAATCCGCCCTGGCGGTGTGGGCCCATCGTGTGTCTTCCCGTCTCTG	1661
Qy	CAATGGAGCTGGTGTCTGTGACCGGTTCGTGTGGGGGGCTGCCCCCTGCCCCCGGTCAA	791
Db	CAATGSGACTGGTGTCTGTGACCGGTTCGTGTGGGGGGCTGCCCCCTGCCCCCGGTCAA	1721
Qy	GTTTGTGGTAGGTGCGAATGACAGCGCCAAACACCTTGACTGGCAATTCCTGTGGCCCTTCT	851
Db	GTTTGTGGTAGGTGCGAATGACAGCGCCAAACACCTTGACTGGCAATTCCTGTGGCCCTTCT	1781
Qy	AGCCTCAGTCCATCCCTTGACCTTTGTGTCAAGTCTCTGAGTACGGAGCCCGCTGGGAGAC	911
Db	AGCCTCAGTCCATCCCTTGACCTTTGTGTCAAGTCTCTGAGTACGGAGCCCGCTGGGAGAC	1841
Qy	GGGCTTAGGCTCCGGGCACTGGCTTCCTGGCAAGTACTTGGGTGAGAGGCAATCGTGCT	971
Db	GGGCTTAGGCTCCGGGCACTGGCTTCCTGGCAAGTACTTGGGTGAGAGGCAATCGTGCT	1901
Qy	GCTGCTCAGTCTGGCCGACGTGACGTGACGAGCTCTCGTCTCCTGTCTCGGGGCTATGG	1031
Db	GCTGCTCAGTCTGGCCGACGTGACGTGACGAGCTCTCGTCTCCTGTCTCGGGGCTATGG	1961
Qy	GAAGTCCCCCTCCCTGGGCAAGCTTCAGAGCAGGGGCTCTAAGGCTGCCCTGGCACTGGCAGG	1091
Db	GAAGTCCCCCTCCCTGGGCAAGCTTCAGAGCAGGGGCTCTAAGGCTGCCCTGGCACTGGCAGG	2021
Qy	GCTGGCCGCGCACTGCCCCCTGGCCTCAGTGGGAGAATA CGGGGCTCCCCCACTTGCCT	1151
Db	GCTGGCCGCGCACTGCCCCCTGGCCTCAGTGGGAGAATA CGGGGCTCCCCCACTTGCCT	2081
Qy	GCCCTACGCGCCACCTGAGGGTACGACAGAGCCCTGGGCTTACCGTGGCCCTGTGTGAT	1211
Db	GCCCTACGCGCCACCTGAGGGTACGACAGAGCCCTGGGCTTACCGTGGCCCTGTGTGAT	2141
Qy	GATGAATCTCTTCTGTGTTCTGTGGTCGTGGCCGTCCTACATCAAACTGTACTGTGACCT	1271
Db	GATGAATCTCTTCTGTGTTCTGTGGTCGTGGCCGTCCTACATCAAACTGTACTGTGACCT	2201
Qy	GCCGCGGGGCACTTTGAGCCGCTGTGGGACTGCGCCATGGTAGGCACTGGCTTGCGCT	1331
Db	GCCGCGGGGCACTTTGAGCCGCTGTGGGACTGCGCCATGGTAGGCACTGGCTTGCGCT	2361
Qy	CATCTTCGACAGCGGGCTCTCTACTGTCCGTGGCCCTTCTCAGCTTCGCTCCATGCT	1391
Db	CATCTTCGACAGCGGGCTCTCTACTGTCCGTGGCCCTTCTCAGCTTCGCTCCATGCT	2321
Qy	GGGCTCTTCCCTGTACGCGCCGAGGCGCTCAAGTCTGTCTGTGGTGTGTGTCGCCCT	1451
Db	GGGCTCTTCCCTGTACGCGCCGAGGCGCTCAAGTCTGTCTGTGGTGTGTGTCGCCCT	2381
Qy	GCCTGCTGCTCAACCCACTGCTGTACTGCTTCAACCCCCACCTTCGGGATGACCT	1511
Db	GCCTGCTGCTCAACCCACTGCTGTACTGCTTCAACCCCCACCTTCGGGATGACCT	2441
Qy	TCGGCGGCTTCGGCCCCGCGAGGGAGCTACAGGGCCCCCTAGCTATGCTCGGGCCGGAGA	1571
Db	TCGGCGGCTTCGGCCCCGCGAGGGAGCTACAGGGCCCCCTAGCTATGCTCGGGCCGGAGA	2501
Qy	GCTGGAGAAGAGCTCTGTGATCTACCCAGGGCCCTGTAGCTTCTCATGTGGATCT	1631
Db	GCTGGAGAAGAGCTCTGTGATCTACCCAGGGCCCTGTAGCTTCTCATGTGGATCT	2561
Qy	CATTCTGGAAAGCTTCTGAAGCTGGGCGGCCCTCTGGGCTGGAGACCTATGGCTTCCCTC	1691
Db	CATTCTGGAAAGCTTCTGAAGCTGGGCGGCCCTCTGGGCTGGAGACCTATGGCTTCCCTC	2621
Qy	AGTGACCTCATCTCTGTGACGAGCAGGGGGCCCCCAGGCTGGAGGGCAGCCATTGTGT	1751
Db	AGTGACCTCATCTCTGTGACGAGCAGGGGGCCCCCAGGCTGGAGGGCAGCCATTGTGT	2681

QY	1752	AGAGCCAGAGGGGNAACACCTTTTGGGAACCCCCAACCTTCATGGATGGAGAACTGCTGCT	1811
Db	2682	AGAGCCAGAGGGGNAACACCTTTTGGGAACCCCCAACCTTCATGGATGGAGAACTGCTGCT	2741
QY	1812	GAGGGCAGAGGGGATCTACGCCAGCAGGTGGAGGCTTGTTCAGGGGGGTGGCGGCTTTTCAGCC	1871
Db	2742	GAGGGCAGAGGGGATCTACGCCAGCAGGTGGAGGCTTGTTCAGGGGGGTGGCGGCTTTTCAGCC	2801
QY	1872	CTCTGGCTTGGCCTTTGCTTTCACACGTG	1899
Db	2802	CTCTGGCTTGGCCTTTGCTTTCACACGTG	2829

```

RESULT 12
US-10-176-847-89
; Sequence 89, Application US/10176847
; Publication No. US20030068636A1
; GENERAL INFORMATION:
; APPLICANT: Velby, Petter Ole
; TITLE OF INVENTION: COMPOSITIONS, KITS, AND METHODS FOR
; TITLE OF INVENTION: IDENTIFICATION, ASSESSMENT, PREVENTION, AND THERAPY OF
; TITLE OF INVENTION: AND OVARIAN CANCER
; FILE REFERENCE: MRI-039
; CURRENT APPLICATION NUMBER: US/10/176,847
; CURRENT FILING DATE: 2002-06-21
; NUMBER OF SEQ ID NOS: 112
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 89
; LENGTH: 3438
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-176-847-89

```

Query Match	92.4%	Score 1754.4	DB 14	Length 3438
Best Local Similarity	99.9%	Pred. No. 0		
Matches 1766	Conservative 0	Mismatches 1	Indels 1	Gaps 1
QY	132	AGGGCCCTCAACACACACACCGCATCTGGGAATTTGGAGCTNGACACCTTCAGCCAGCTGAG	191	
Db	1187	ATGGCCCTCAACACACACACCGCATCTGGGAATTTGGAGCT--GACACCTTCAGCCAGCTGAG	1245	
QY	192	CTCCCTGCAAGCCCTGGATCTTAGCTGGAAAGCGCATCCGGTCCATCCACCTGAGGCCTT	251	
Db	1246	CTCCCTGCAAGCCCTGGATCTTAGCTGGAAAGCGCATCCGGTCCATCCACCTGAGGCCTT	1305	
QY	252	CTCCACCCCTGGACTCCCTCTGTGTGAAGCTGGACCTGACAGACAACAGGTGACACACACTGCC	311	
Db	1306	CTCCACCCCTGGACTCCCTCTGTGTGAAGCTGGACCTGACAGACAACAGGTGACACACACTGCC	1365	
QY	312	CTTGGCTGGACTTGGGGGCTTGATGATCTGAAGCTCAAAAGGAACTTGCTCTCTCCCA	371	
Db	1366	CTTGGCTGGACTTGGGGGCTTGATGATCTGAAGCTCAAAAGGAACTTGCTCTCTCCCA	1425	
QY	372	GGCCTTCTCCAAGGACAGATTTCCTCAAACTGAGGATCTCTGGAGGTGCTTATGCCTACCA	431	
Db	1426	GGCCTTCTCCAAGGACAGATTTCCTCAAACTGAGGATCTCTGGAGGTGCTTATGCCTACCA	1485	
QY	432	GTGCTGTCCCTATGGGATGTGTGCGCAGTTCCTTCAAGGCCTCTTGGGAGTGGGAGGCTGA	491	
Db	1486	GTGCTGTCCCTATGGGATGTGTGCGCAGTTCCTTCAAGGCCTCTTGGGAGTGGGAGGCTGA	1545	
QY	492	AGACCTTCACCTTGATGATGAGAGTGCTTCAAAAGGCCCTCTGGGCTCTCTTCCAGACA	551	
Db	1546	AGACCTTCACCTTGATGATGAGAGTGCTTCAAAAGGCCCTCTGGGCTCTCTTCCAGACA	1605	
QY	552	AGCAGAGAACCACTATGACACGAGCACTTGGATGAGCTCCAGCTGGAGATGGAGACTCAAA	611	
Db	1606	AGCAGAGAACCACTATGACACGAGCACTTGGATGAGCTCCAGCTGGAGATGGAGACTCAAA	1665	
QY	612	GCCACACCCGAGTGTCCAGTGTAGCCCTACTCCAGGCCCTCTTCAAGCCCTGTGAGTACCT	671	
Db	1666	GCCACACCCGAGTGTCCAGTGTAGCCCTACTCCAGGCCCTCTTCAAGCCCTGTGAGTACCT	1725	

Qy	672	CTTTGAAAGCTGGGGCATCCGCCCTGGCGTGTGGGCCATCGTTGCTCTCCCGTCTCTG	731
Db	1726	CTTTGAAAGCTGGGGCATCCGCCCTGGCGTGTGGGCCATCGTTGCTCTCCCGTCTCTG	1785
Qy	732	CAATGGACGTGTTGCTGTGACCGTGTTCGCTGGCGGGCTGCCCCCTGCCCGGTCAA	791
Db	1786	CAATGGACGTGTTGCTGTGACCGTGTTCGCTGGCGGGCTGCCCCCTGCCCGGTCAA	1845
Qy	792	GTTTGTGTAGTGCATTTGGATTGCAGCGCCAAACACTTGACTGGCATTTCCCTGTGGCCTTCT	851
Db	1846	GTTTGTGTAGTGCATTTGGATTGCAGCGCCAAACACTTGACTGGCATTTCCCTGTGGCCTTCT	1905
Qy	852	AGCCTCAGTCGATCCCTCGACCTTTTGTGTAAGTTCTCTAGTACCGAGCCCGCTGGGAGAC	911
Db	1906	AGCCTCAGTCGATCCCTCGACCTTTTGTGTAAGTTCTCTAGTACCGAGCCCGCTGGGAGAC	1965
Qy	912	GGGGCTAGGCTGCCGGGCCACTGGCTTCCTGGCAGTACTTGGGTGCGAGGCATCGGTGCT	971
Db	1966	GGGGCTAGGCTGCCGGGCCACTGGCTTCCTGGCAGTACTTGGGTGCGAGGCATCGGTGCT	2025
Qy	972	GCTGCTCACTCTGGCCGAGTCAGTCGAGCGTCTCCGTCTCCCTGTCTCGGGCTATGG	1031
Db	2026	GCTGCTCACTCTGGCCGAGTCAGTCGAGCGTCTCCGTCTCCCTGTCTCGGGCTATGG	2085
Qy	1032	GAATGCCCTCCCTGGGCAGCGTTTCAGCAGAGGGTCTTAGCTGCTCGTGCACCTGGCAGG	1091
Db	2086	GAATGCCCTCCCTGGGCAGCGTTTCAGCAGAGGGTCTTAGCTGCTCGTGCACCTGGCAGG	2145
Qy	1092	GCTGGCGCGCACTGCCCTGGCCTCAGTGGGAGAAATAGGGGCGCTCCCACTCTGCCT	1151
Db	2146	GCTGGCGCGCACTGCCCTGGCCTCAGTGGGAGAAATAGGGGCGCTCCCACTCTGCCT	2205
Qy	1152	GCCCTACGCGCCACTCAGGGGTACGACAGAGCCCTGGGGTTACCGTGGCCCTGGTGAT	1211
Db	2206	GCCCTACGCGCCACTCAGGGGTACGACAGAGCCCTGGGGTTACCGTGGCCCTGGTGAT	2265
Qy	1212	GATGAACCTCTTCTGTTTCTGTGTCGTGGCCGCGTCAATCAAACTGTACTGTGACCT	1271
Db	2266	GATGAACCTCTTCTGTTTCTGTGTCGTGGCCGCGTCAATCAAACTGTACTGTGACCT	2325
Qy	1272	GCCGCGGGCGACTTTCAGGCCGTGTGCGACTGCGCATGATGTAGGCACAGTGGCCTGGCT	1331
Db	2326	GCCGCGGGCGACTTTCAGGCCGTGTGCGACTGCGCATGATGTAGGCACAGTGGCCTGGCT	2385
Qy	1332	CATCTTCGACAGCGGGTCTCTACTGTCCGTGGCCCTTCTCAGCTTCGCCCTCCATGCT	1391
Db	2386	CATCTTCGACAGCGGGTCTCTACTGTCCGTGGCCCTTCTCAGCTTCGCCCTCCATGCT	2445
Qy	1392	GGGGCTCTTTCGCTGTACGCCCGAGGCGGTCAAGTCTGTCTGCTGTGTGTGTCGCCCT	1451
Db	2446	GGGGCTCTTTCGCTGTACGCCCGAGGCGGTCAAGTCTGTCTGCTGTGTGTGTCGCCCT	2505
Qy	1452	GCTGCTGCTCAACCCACTGCTGTACTGTCTTCAACCCCCACTTCGGGATGACCT	1511
Db	2506	GCTGCTGCTCAACCCACTGCTGTACTGTCTTCAACCCCCACTTCGGGATGACCT	2565
Qy	1512	TCGGCGCTTCGGCCCCCGCAGGGGACTCAGGGCCCTTAGCCTTAGCTCGGCCCGGGGA	1571
Db	2566	TCGGCGCTTCGGCCCCCGCAGGGGACTCAGGGCCCTTAGCCTTAGCTCGGCCCGGGGA	2625
Qy	1572	GCTGGAGAGCTCCTGTGATTTCTACCCAGGCCCTGGTAGCCTTCTCATGTGGATCT	1631
Db	2626	GCTGGAGAGCTCCTGTGATTTCTACCCAGGCCCTGGTAGCCTTCTCATGTGGATCT	2685
Qy	1632	CATTCTGGAAAGCTTCTGAAGCTGGGGCGGCCCTCTGGGCTGGAGACCTATGGCTTCCGCTC	1691
Db	2686	CATTCTGGAAAGCTTCTGAAGCTGGGGCGGCCCTCTGGGCTGGAGACCTATGGCTTCCGCTC	2745
Qy	1692	AGTGAACCTCATCTCTGTTCAGACAGCAGGGGGCCCCCAGGCTGGAGGACCAATGTGT	1751
Db	2746	AGTGAACCTCATCTCTGTTCAGACAGCAGGGGGCCCCCAGGCTGGAGGACCAATGTGT	2805

Qy	1752	AGAGCCAGAGGGGAAACACATTTTGGGAACCCCCAACCTTCCATGGATGGAGAACTGCTGCT	1811
Db	2806	AGAGCCAGAGGGGAAACACATTTTGGGAACCCCCAACCTTCCATGGATGGAGAACTGCTGCT	2865
Qy	1812	GAGGGCAGAGGGATCTACGCCACAGGTGAGAGCTTGTTCAGGGGGTGGCGGCTTTTCAGCC	1871
Db	2866	GAGGGCAGAGGGATCTACGCCACAGGTGAGAGCTTGTTCAGGGGGTGGCGGCTTTTCAGCC	2925
Qy	1872	CTCTGGCTTGGCCCTTGTGCTTTCACACGTG	1899
Db	2926	CTCTGGCTTGGCCCTTGTGCTTTCACACGTG	2953
RESULT 13			
US-10-331-496A-94			
; Sequence 94, Application US/10331496A			
; Publication No. US20030228305A1			
; GENERAL INFORMATION:			
; APPLICANT: FRANTZ, GRETCHEN			
; APPLICANT: HILLAN, KENNETH J.			
; APPLICANT: PHILLIPS, HEIDI S.			
; APPLICANT: POLAKIS, PAUL			
; APPLICANT: SMITH, VICTORIA			
; APPLICANT: SPENCER, SUSAN D.			
; APPLICANT: WILLIAMS, P. MICKEY			
; APPLICANT: WU, THOMAS D.			
; APPLICANT: ZHANG, ZEMIN			
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE DIAGNOSIS AND			
; TREATMENT OF TUMOR			
; FILE REFERENCE: P5014R1-PCT			
; CURRENT APPLICATION NUMBER: US/10/331,496A			
; CURRENT FILING DATE: 2002-12-30			
; PRIOR APPLICATION NUMBER: US 60/345,444			
; PRIOR FILING DATE: 2002-01-02			
; PRIOR APPLICATION NUMBER: US 60/351,885			
; PRIOR FILING DATE: 2002-01-25			
; PRIOR APPLICATION NUMBER: US 60/360,066			
; PRIOR FILING DATE: 2002-02-25			
; PRIOR APPLICATION NUMBER: US 60/362,004			
; PRIOR FILING DATE: 2002-03-05			
; PRIOR APPLICATION NUMBER: US 60/366,869			
; PRIOR FILING DATE: 2002-03-20			
; PRIOR APPLICATION NUMBER: US 60/366,284			
; PRIOR FILING DATE: 2002-03-21			
; PRIOR APPLICATION NUMBER: US 60/368,679			
; PRIOR FILING DATE: 2002-03-28			
; PRIOR APPLICATION NUMBER: US 60/404,809			
; PRIOR FILING DATE: 2002-08-19			
; PRIOR APPLICATION NUMBER: US 60/405,645			
; PRIOR FILING DATE: 2002-08-21			
; NUMBER OF SEQ ID NOS: 95			
; SEQ ID NO 94			
; LENGTH: 3443			
; TYPE: DNA			
; ORGANISM: Homo sapien			
US-10-331-496A-94			
Query Match 92.4%; Score 1754.4; DB 17; Length 3443;			
Best Local Similarity 99.9%; Pred.No. 0;			
Matches 1766; Conservative 0; Mismatches 1; Indels 1; Gaps 1;			
Qy	132	AGCGGCTCCACACAAACCGCATCTGGGAATTTGGAGCTTGACACCTTCAGCCAGCTGAG	191
Db	1189	ATCGGCTCCACACAAACCGCATCTGGGAATTTGGAGCTTGACACCTTCAGCCAGCTGAG	1247
Qy	192	CTCCCTGCAAGCCCTTGATCTTTAGCTGGAAACGCCATCCCGTCCATCCACCCCTGAGGCCTT	251
Db	1248	CTCCCTGCAAGCCCTTGATCTTTAGCTGGAAACGCCATCCCGTCCATCCACCCCTGAGGCCTT	1307
Qy	252	CTCCACCTTGCACTCCCTGGTCAAGCTGGACCTGACAGACAACACCGACTGACCACTGCC	311
Db	1308	CTCCACCTTGCACTCCCTGGTCAAGCTGGACCTGACAGACAACACCGACTGACCACTGCC	1367

Qy	312	CTGGCTGGACTTGGGGCTTGATGCACTCTGAAGCTCAAAGGAACTTGTCTCTCTCCCA	371
Db	1368	CTGGCTGGACTTGGGGCTTGATGCACTCTGAAGCTCAAAGGAACTTGTCTCTCTCCCA	1427
Qy	372	GGCCTTCTCCAAGGACAGTGTTCACAAACTCAGAGATCTCGAGAGTGCTTATGCGCTACCA	431
Db	1428	GGCCTTCTCCAAGGACAGTGTTCACAAACTCAGAGATCTCGAGAGTGCTTATGCGCTACCA	1487
Qy	432	GTGCTGTCCCTPATGGGATGTGTGCCAGCTTCTTCAAGGCCCTCTGGGCAGTGGGAGGCTGA	491
Db	1488	GTGCTGTCCCTPATGGGATGTGTGCCAGCTTCTTCAAGGCCCTCTGGGCAGTGGGAGGCTGA	1547
Qy	492	AGACCTTCACCTTGATGATGAGGAGTCTTCAAAAAGGCCCTCTGGGCCCTCTTGTGCCAGACA	551
Db	1548	AGACCTTCACCTTGATGATGAGGAGTCTTCAAAAAGGCCCTCTGGGCCCTCTTGTGCCAGACA	1607
Qy	552	AGCAGAGAACCACTATGACCAAGGACCTGGATAGCTCCAGCTGAGATGGAGGACTCAAA	611
Db	1608	AGCAGAGAACCACTATGACCAAGGACCTGGATAGCTCCAGCTGAGATGGAGGACTCAAA	1667
Qy	612	GCCACACCCAGTGTCCAGTGTAGCCCTACTTCCAGGGCCCTTCAAGGCCCTGTGAGTACCT	671
Db	1668	GCCACACCCAGTGTCCAGTGTAGCCCTACTTCCAGGGCCCTTCAAGGCCCTGTGAGTACCT	1727
Qy	672	CTTTGAAAGCTGGGGCATCCGCCCTGGCGCTGTGGGCCAATCGTGTGTCTCTCGTGTCTG	731
Db	1728	CTTTGAAAGCTGGGGCATCCGCCCTGGCGCTGTGGGCCAATCGTGTGTCTCTCGTGTCTG	1787
Qy	732	CAATGGACTGTGTCTGTGACCGGTTCCTGGCGGGCCCTGCCCCCTGCCCGGTCAA	791
Db	1788	CAATGGACTGTGTCTGTGACCGGTTCCTGGCGGGCCCTGCCCCCTGCCCGGTCAA	1847
Qy	792	GTTTGTGGTAGGTTCGGATTGACGGCGCCAAACACCTTGACTGGCAATTTCTGTGGCCTTCT	851
Db	1848	GTTTGTGGTAGGTTCGGATTGACGGCGCCAAACACCTTGACTGGCAATTTCTGTGGCCTTCT	1907
Qy	852	AGCCTCAGTCGATGCCCTGACCTTTGGTTCAGTTCTCTGAGTACGGAGCCCGCTGGGAGAC	911
Db	1908	AGCCTCAGTCGATGCCCTGACCTTTGGTTCAGTTCTCTGAGTACGGAGCCCGCTGGGAGAC	1967
Qy	912	GGGGCTAGGCTCCGGGGCCACTGGACTTCTCGGCACTATTGGGTTCGAGGCATCGGTCT	971
Db	1968	GGGGCTAGGCTCCGGGGCCACTGGACTTCTCGGCACTATTGGGTTCGAGGCATCGGTCT	2027
Qy	972	GCTGCTCACTCTGGCCGAGTGACGTGCAGCGTCTCCGTCTCTCTGTGTCGGGGCTATGG	1031
Db	2028	GCTGCTCACTCTGGCCGAGTGACGTGCAGCGTCTCCGTCTCTCTGTGTCGGGGCTATGG	2087
Qy	1032	GAAGTCCCTCTCTGCGGCACGTTTCAGACAGGGGTCTTAGGTGCTGCTGCACTGGCAGG	1091
Db	2088	GAAGTCCCTCTCTGCGGCACGTTTCAGACAGGGGTCTTAGGTGCTGCTGCACTGGCAGG	2147
Qy	1092	GCTGGCCGCGCACTGCCCTTGGCTCAGTGGGAGAAATACGGGGCTCTCCCACTCTGCCT	1151
Db	2148	GCTGGCCGCGCACTGCCCTTGGCTCAGTGGGAGAAATACGGGGCTCTCCCACTCTGCCT	2207
Qy	1152	GCCCTACGCGCCACTGAGGGTTCAGCAGAGCCCTGGGCTTACCGTGCCTGCTGGTGAT	1211
Db	2208	GCCCTACGCGCCACTGAGGGTTCAGCAGAGCCCTGGGCTTACCGTGCCTGCTGGTGAT	2267
Qy	1212	GATGAACTCCTTCTGTTTCTCTGGTTCGTGGCGGTGCCCTACATCAAACTGTACTGTGACCT	1271
Db	2268	GATGAACTCCTTCTGTTTCTCTGGTTCGTGGCGGTGCCCTACATCAAACTGTACTGTGACCT	2327
Qy	1272	GCCCGGGGGCACTTTAGGCGGTGTGGGACTGGGCCATAGGTGAGGCACTGGCGCTGGCT	1331
Db	2328	GCCCGGGGGCACTTTAGGCGGTGTGGGACTGGGCCATAGGTGAGGCACTGGCGCTGGCT	2387
Qy	1332	CATCTTCGCAGACGGGCTCTCTACTGTCCGTGGCCCTTCTCAGCTTCGCCCTCCATGCT	1391
Db	2388	CATCTTCGCAGACGGGCTCTCTACTGTCCGTGGCCCTTCTCAGCTTCGCCCTCCATGCT	2447
Qy	1392	GGGCTCTTCCCTGTCAAGCCGAGGCGCGTCAAGTCTGTCTGTGTGGTGTGTGCGCCCT	1451

Db	2448	GGGCGCTCTTCCCTGTCTACGCGCCGAGGCGGTCAAGTCTGCTCGTGGTGTGTCGCCCT	2507
Qy	1452	GCTGCGCTGCTCAACCCACTGTGTACCTGCTCTTCAACCCCACTTCCGGGATGACCT	1511
Db	2508	GCTGCGCTGCTCAACCCACTGTGTACCTGCTCTTCAACCCCACTTCCGGGATGACCT	2567
Qy	1512	TCGCGCGCTTCGGGCCCGCGACGAGGAGCTCAGGGCCCCCTAGCTATGCTGCGGCCGGGGA	1571
Db	2568	TCGCGCGCTTCGGGCCCGCGACGAGGAGCTCAGGGCCCCCTAGCTATGCTGCGGCCGGGGA	2627
Qy	1572	GCTGGAGAAAGCTCCTGTGATTTACTACCAAGGCGCTGGTAGCCTTCTCTGATGTGATCT	1631
Db	2628	GCTGGAGAAAGCTCCTGTGATTTACTACCAAGGCGCTGGTAGCCTTCTCTGATGTGATCT	2687
Qy	1632	CATTCTGGAAGCTTCTGAACTGGGGGGCCCCCTGGGCTGGAGACCTATGGCTTCCCCCTC	1691
Db	2688	CATTCTGGAAGCTTCTGAACTGGGGGGCCCCCTGGGCTGGAGACCTATGGCTTCCCCCTC	2747
Qy	1692	AGTGACCTCATCTCTGTGTAGCAGCAGAGGGGCCCAAGCTGGAGGGGACGCAATTGTGT	1751
Db	2748	AGTGACCTCATCTCTGTGTAGCAGCAGAGGGGCCCAAGCTGGAGGGGACGCAATTGTGT	2807
Qy	1752	AGAGCCAGAGGGGAAACCACTTTGGGAAACCCCAACCTCCATGGATGGAGAACTGTGCT	1811
Db	2808	AGAGCCAGAGGGGAAACCACTTTGGGAAACCCCAACCTCCATGGATGGAGAACTGTGCT	2867
Qy	1812	GAGGGCAGAGGGGATCTACGCCACGAGGTGAGAGCTTGTTCAGGGGGGTGGCGGCTTTCAGCC	1871
Db	2868	GAGGGCAGAGGGGATCTACGCCACGAGGTGAGAGCTTGTTCAGGGGGGTGGCGGCTTTCAGCC	2927
Qy	1872	CTCTGGCTTGGCCTTTGGCTTTCACAGTG	1899
Db	2928	CTCTGGCTTGGCCTTTGGCTTTCACAGTG	2955

RESULT 14

```

RES001 14
US-09-851-595-10
; Sequence 10, Application US/09851595
; Publication No. US20030166047A1
; GENERAL INFORMATION:
; APPLICANT: Gu, Wei
; TITLE OF INVENTION: NOVEL G-PROTEIN COUPLED RECEPTORS AND USES THEREFOR
; FILE REFERENCE: WNI-080CP
; CURRENT APPLICATION NUMBER: US/09/851,595
; CURRENT FILING DATE: 2000-05-08
; PRIOR APPLICATION NUMBER: 09/556,588
; PRIOR FILING DATE: 2000-05-08
; PRIOR APPLICATION NUMBER: 60/132,896
; PRIOR FILING DATE: 1999-05-06
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 10
; LENGTH: 3492
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (104)...(3004)
US-09-851-595-10

```

Query Match 92.4%; Score 1754.4; DB 10; Length 3492;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1766; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

Qy	132	AGGGCTCAACACAAACCGCATCTGGAAATTTGGAGCTTGACCTTTAGCCAGCTGAG	191
Db	1238	ATGGCTCTCAACACAAACCGCATCTGGAAATTTGGAGCT-GACACCTTTAGCCAGCTGAG	1286
Qy	192	CTCCCTCGAAGCCCTGGATCTTAGCTGGAAAGCCCATCCGGTCCATCCACCTTGAGGCCTT	251
Db	1297	CTCCCTCGAAGCCCTGGATCTTAGCTGGAAAGCCCATCCGGTCCATCCACCTTGAGGCCTT	1356

Qy 252 CTCACACCTGACATCCCTGGTCAAGCTGGAGCTTGACAGACAAACAGCTGACCACTGCC 311
Db 1357 CTCACACCTGACATCCCTGGTCAAGCTGGAGCTTGACAGACAAACAGCTGACCACTGCC 1416
Qy 312 CCTGGCTGGACTTGGGGCTTGTAGTCATCTGAAGCTCAAAGGAACTTGTCTCTCCCA 371
Db 1417 CCTGGCTGGACTTGGGGCTTGTAGTCATCTGAAGCTCAAAGGAACTTGTCTCTCCCA 1476
Qy 372 GGCCTTCTCCAGGACAGTTTCCCAAACTGAGGATCTCGAGGTGCTTATGCTTACCA 431
Db 1477 GGCCTTCTCCAGGACAGTTTCCCAAACTGAGGATCTCGAGGTGCTTATGCTTACCA 1536
Qy 432 GTGCTGCTCCCTATGGGATGTGTCAGCTTCTTCAAGGCCCTCTGGGAGTGGAGGTGA 491
Db 1537 GTGCTGCTCCCTATGGGATGTGTCAGCTTCTTCAAGGCCCTCTGGGAGTGGAGGTGA 1596
Qy 492 AGACCTTTCACCTTGATGATGAGGAGTCTTCAAAAAGGCCCTTGGGCCCTCTTGGCAGACA 551
Db 1597 AGACCTTTCACCTTGATGATGAGGAGTCTTCAAAAAGGCCCTTGGGCCCTCTTGGCAGACA 1656
Qy 552 AGCAGAGAACACTATGACACGAGGACCTGGATGAGCTCCAGCTGGAGATGGAGGACTCAA 611
Db 1657 AGCAGAGAACACTATGACACGAGGACCTGGATGAGCTCCAGCTGGAGATGGAGGACTCAA 1716
Qy 612 GCCACACCCAGTGTCCAGTGTAGCCCTACTCCAGGCCCTTCAAGGCCCTGTGAGTACCT 671
Db 1717 GCCACACCCAGTGTCCAGTGTAGCCCTACTCCAGGCCCTTCAAGGCCCTGTGAGTACCT 1776
Qy 672 CTTTGAAGCTGGGGCATCCGCCCTGGCGGTGTGGGCCATCGTTGTCTCCGTGCTCTG 731
Db 1777 CTTTGAAGCTGGGGCATCCGCCCTGGCGGTGTGGGCCATCGTTGTCTCCGTGCTCTG 1836
Qy 732 CAATGGAGCTGTGCTGCTGACCGTGTTCGCTGGCGGGCTGCCGCCCTGCCCGGTCAA 791
Db 1837 CAATGGAGCTGTGCTGCTGACCGTGTTCGCTGGCGGGCTGCCGCCCTGCCCGGTCAA 1896
Qy 792 GTTTGTGTAGTGTGGATGTGAGCGCGCAACACCTTGACTGGCATTTCTGTGCGCTTCT 851
Db 1897 GTTTGTGTAGTGTGGATGTGAGCGCGCAACACCTTGACTGGCATTTCTGTGCGCTTCT 1956
Qy 852 AGCCTCAGTCGATGCCCTGACCTTTTGGTTCAGTTCTCTGAGTACGAGCGCCGCTGGAGAC 911
Db 1957 AGCCTCAGTCGATGCCCTGACCTTTTGGTTCAGTTCTCTGAGTACGAGCGCCGCTGGAGAC 2016
Qy 912 GGGGCTAGGCTGCCGGGCACTGCTTCTGGCAGTACTGGGTGCGAGGACATCGGTGCT 971
Db 2017 GGGGCTAGGCTGCCGGGCACTGCTTCTGGCAGTACTGGGTGCGAGGACATCGGTGCT 2076
Qy 972 GCTGCTCACTCTGCGCGCAGTGACGTGCGGTCTCCGTCTCTGTGTGTCGGGCTATGG 1031
Db 2077 GCTGCTCACTCTGCGCGCAGTGACGTGCGGTCTCCGTCTCTGTGTGTCGGGCTATGG 2136
Qy 1032 GAAGTCCCTTCTCTGGGCACTGCTTTCGAGCAGGGGTCTTAGGCTCCCTGGCACTGGCAGG 1091
Db 2137 GAAGTCCCTTCTCTGGGCACTGCTTTCGAGCAGGGGTCTTAGGCTCCCTGGCACTGGCAGG 2196
Qy 1092 GCTGGCGCGCACTGCCCTGGCTCAGTGGGAGAAATACGGGCCCTCCCACTCTGCCT 1151
Db 2197 GCTGGCGCGCACTGCCCTGGCTCAGTGGGAGAAATACGGGCCCTCCCACTCTGCCT 2256
Qy 1152 GCCCTACGCGCACTCAGGGTACGAGCAGCGCCCTGGGCTTACCGTGGCCCTGGTGAT 1211
Db 2257 GCCCTACGCGCACTCAGGGTACGAGCAGCGCCCTGGGCTTACCGTGGCCCTGGTGAT 2316
Qy 1212 GATGAATCTCTTCTGTTTCTGTGCTCGTGGCGGTGCCCTACATCAAACTGTACTGTGACCT 1271
Db 2317 GATGAATCTCTTCTGTTTCTGTGCTCGTGGCGGTGCCCTACATCAAACTGTACTGTGACCT 2376
Qy 1272 GCOCGGGGGCACTTTAGGGCGGTGTGGAGCTGGCCATGCTGAGGACGCTGGCTGGCT 1331
Db 2377 GCOCGGGGGCACTTTAGGGCGGTGTGGAGCTGGCCATGCTGAGGACGCTGGCTGGCT 2436
Qy 1332 CATCTTCGAGAGGGGCTCTCTACTGTCCCGTGGCCCTTCTCAGCTTCGCTCCATGCT 1391

Db 2437 CATCTTCGAGAGGGGCTCTCTACTGTCCCGTGGCCCTTCTCAGCTTCGCTCCATGCT 2496
Qy 1392 GGGCCTCTTCCCTGTGACGGCCGAGGGCGTCAAGTCTGTCTGTGCTGTGCTGCCCCCT 1451
Db 2497 GGGCCTCTTCCCTGTGACGGCCGAGGGCGTCAAGTCTGTCTGTGCTGTGCTGCCCCCT 2556
Qy 1452 GCCTGCCTGCTCAACCCACTGTGTACTGCTTCAACCCCCACTTCCGGGATGACCT 1511
Db 2557 GCCTGCCTGCTCAACCCACTGTGTACTGCTTCAACCCCCACTTCCGGGATGACCT 2616
Qy 1512 TCGGCGGCTTCGGGCCCGCGCAGGGGACTCAGGGGCCCTAGCCCTATGCTGCGGCCGGGA 1571
Db 2617 TCGGCGGCTTCGGGCCCGCGCAGGGGACTCAGGGGCCCTAGCCCTATGCTGCGGCCGGGA 2676
Qy 1572 GCTGGAGAGAGCTCTCTGTGATTTACCCAGGCCCTCGTAGCCTTCTCTGATGTGGATCT 1631
Db 2677 GCTGGAGAGAGCTCTCTGTGATTTACCCAGGCCCTCGTAGCCTTCTCTGATGTGGATCT 2736
Qy 1632 CATTCGGAAGCTTCTGAAGCTGGGGGCCCTTGGGCTGGAGACCTATGGCTTCCCTC 1691
Db 2737 CATTCGGAAGCTTCTGAAGCTGGGGGCCCTTGGGCTGGAGACCTATGGCTTCCCTC 2796
Qy 1692 AGTGACCTCATCTCTCTGTGACGAGCCAGGGGCCCTCCAGGCTGGAGGCGAGCCATTGTGT 1751
Db 2797 AGTGACCTCATCTCTCTGTGACGAGCCAGGGGCCCTCCAGGCTGGAGGCGAGCCATTGTGT 2856
Qy 1752 AGAGCCAGAGGGGAAACACATTTGGGAAACCCCCAAACCTCATGATGAGAGAACTGTGCT 1811
Db 2857 AGAGCCAGAGGGGAAACACATTTGGGAAACCCCCAAACCTCATGATGAGAGAACTGTGCT 2916
Qy 1812 GAGGCGAGAGGGATCTACGCCAGCGTGGAGGCTTGTCTAGGGGTGCGGCTTTCAGCC 1871
Db 2917 GAGGCGAGAGGGATCTACGCCAGCGTGGAGGCTTGTCTAGGGGTGCGGCTTTCAGCC 2976
Qy 1872 CTCTGGCTTGGCCCTTTGCTTTCACACGTG 1899
Db 2977 CTCTGGCTTGGCCCTTTGCTTTCACACGTG 3004

RESULT 15

US-10-664-667-10
; Sequence 10, Application US/10664667
; Publication No. US20040058377A1
; GENERAL INFORMATION:
; APPLICANT: Gu, Wei
; TITLE OF INVENTION: NOVEL G-PROTEIN COUPLED RECEPTORS AND USES THEREFOR
; FILE REFERENCE: WNI-080CP
; CURRENT APPLICATION NUMBER: US/10/664,667
; PRIOR FILING DATE: 2003-09-18
; PRIOR APPLICATION NUMBER: US/09/851,595
; PRIOR FILING DATE: 2000-05-08
; PRIOR APPLICATION NUMBER: 09/556,588
; PRIOR FILING DATE: 2000-05-08
; PRIOR APPLICATION NUMBER: 60/132,896
; PRIOR FILING DATE: 1999-05-06
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 10
; LENGTH: 3492
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (104) ... (3004)
US-10-664-667-10

Query Match 92.4%; Score 1754.4; DB 17; Length 3492;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1766; Conservative 0; Mismatches 1; Indels 1; Gaps 1;
Qy 132 AGCGGCTCCAAACACACACCGCATCTGGGAAATTTGGAGCTNGACACCTTCAGCCAGCTGAG 191

Db 1238 ATCGGCTTCCAAACAAACCGCATCTGGGAAATTTGGAGCT-GACACCTTTACGCCAGCTGAG 1296
Qy 192 CTCCTGCAAGCCCTGGATCTTTAGCTGGAAGCCATCCGGTCCATCCACCCCTGAGGCCTT 251
Db 1297 CTCCTGCAAGCCCTGGATCTTTAGCTGGAAGCCATCCGGTCCATCCACCCCTGAGGCCTT 1356
Qy 252 CTCACCCCTGCACTCCCTGCTGCTCAAGCTGGAAGCTGACAGAACCAAGCTGACCACTGCC 311
Db 1357 CTCACCCCTGCACTCCCTGCTGCTCAAGCTGGAAGCTGACAGAACCAAGCTGACCACTGCC 1416
Qy 312 CCTGGCTGGAATCTGGGGCTTGATGCATCTGAAGCTCAAGGGAACCTTGCTCTCTCCCA 371
Db 1417 CCTGGCTGGAATCTGGGGCTTGATGCATCTGAAGCTCAAGGGAACCTTGCTCTCTCCCA 1476
Qy 372 GGCTTTCTCCAAAGACAGTTTCCCAAAACTGAGGATCTCGAGGTGCCCTTATGCCATCCA 431
Db 1477 GGCTTTCTCCAAAGACAGTTTCCCAAAACTGAGGATCTCGAGGTGCCCTTATGCCATCCA 1536
Qy 432 GTGCTGCTCCATGAGGATGTGTGCAGCTTCTTCAAGGCCCTCTGGGCAAGTGGGAGGTGA 491
Db 1537 GTGCTGCTCCATGAGGATGTGTGCAGCTTCTTCAAGGCCCTCTGGGCAAGTGGGAGGTGA 1596
Qy 492 AGACCTTCACTGATGATGAGGAGTCTTCAAAAGGCCCTTGGGCTCTTGGCCAGACA 551
Db 1597 AGACCTTCACTGATGATGAGGAGTCTTCAAAAGGCCCTTGGGCTCTTGGCCAGACA 1656
Qy 552 AGCAGAGAACCACTATGACAGGACCTGGATGAGCTCCAGCTGGAGATGGAGGACTCAAA 611
Db 1657 AGCAGAGAACCACTATGACAGGACCTGGATGAGCTCCAGCTGGAGATGGAGGACTCAAA 1716
Qy 612 GCCACACCCCAAGTGTCCAGTGTAGCCCTTATCCAGGCCCTTCAAGGCCCTGTGAGTACCT 671
Db 1717 GCCACACCCCAAGTGTCCAGTGTAGCCCTTATCCAGGCCCTTCAAGGCCCTGTGAGTACCT 1776
Qy 672 CTTTGTGAAGCTGGGGCATCCGCTGGCGTGTGGCCATCGTGTGCTCGGTGCTCTG 731
Db 1777 CTTTGTGAAGCTGGGGCATCCGCTGGCGTGTGGCCATCGTGTGCTCGGTGCTCTG 1836
Qy 732 CAATGGAGCTGTGTGCTGACCGTGTTCGCTGGGGGCTGCGCCCTGCGCCCGGTCAA 791
Db 1837 CAATGGAGCTGTGTGCTGACCGTGTTCGCTGGGGGCTGCGCCCTGCGCCCGGTCAA 1896
Qy 792 GTTTGTGATAGGTGCGATTGACGGCGCCAAACCTTTGACTGCGCATTTCTGTGCGCTTCT 851
Db 1897 GTTTGTGATAGGTGCGATTGACGGCGCCAAACCTTTGACTGCGCATTTCTGTGCGCTTCT 1956
Qy 852 AGCCTCAGTCATGCCCTGACCTTTGGTCACTTCTGAGTACGGAGCGCCCTGGGAGAC 911
Db 1957 AGCCTCAGTCATGCCCTGACCTTTGGTCACTTCTGAGTACGGAGCGCCCTGGGAGAC 2016
Qy 912 GGGGCTAGGCTGCGGGGCACTGGCTTCTTGCGCAGTACTTGGGTGCGAGGCATCGGTGCT 971
Db 2017 GGGGCTAGGCTGCGGGGCACTGGCTTCTTGCGCAGTACTTGGGTGCGAGGCATCGGTGCT 2076
Qy 972 GCTGCTCACTGTGGCCGAGTGCAGTGCAGCGTCTCCGTCTCTGTGTCCGGGCTATGG 1031
Db 2077 GCTGCTCACTGTGGCCGAGTGCAGTGCAGCGTCTCCGTCTCTGTGTCCGGGCTATGG 2136
Qy 1032 GAACTCCCCCTCCCTGGGCGAGGTTTCGACAGGGGTCTTAGGCTGCCCTGGGCACTGGCAGG 1091
Db 2137 GAACTCCCCCTCCCTGGGCGAGGTTTCGACAGGGGTCTTAGGCTGCCCTGGGCACTGGCAGG 2196
Qy 1092 GCTGGCCCGGCACCTGGCCCTGGGCTCAGTGGGGAATACGGGGGCTCCGCCACTCTGCT 1151
Db 2197 GCTGGCCCGGCACCTGGCCCTGGGCTCAGTGGGGAATACGGGGGCTCCGCCACTCTGCT 2256
Qy 1152 GCCCTACGCGCCACCTGAGGGTTCAGCCAGAGCCCTGGGCTTTCACCGTGGCCCTGGTGTAT 1211
Db 2257 GCCCTACGCGCCACCTGAGGGTTCAGCCAGAGCCCTGGGCTTTCACCGTGGCCCTGGTGTAT 2316
Qy 1212 GATGAATCTCTTCTGTTTCTGCTGCTGGCGGTGCTTACATCAAACTGTATGTGACCT 1271
Db 2317 GATGAATCTCTTCTGTTTCTGCTGCTGGCGGTGCTTACATCAAACTGTATGTGACCT 2376

Qy 1272 GCCCGGGGCGAATTTGAGGCCGTGTGGGACTGGCCATGTGAGGCACGTGGCTTGCT 1331
Db 2377 GCCCGGGGCGAATTTGAGGCCGTGTGGGACTGGCCATGTGAGGCACGTGGCTTGCT 2436
Qy 1332 CATCTTGCAGACGGGCTCTTACTGTCCCGTGGCCTTCTCAGCTTCCGCTCCATGCT 1391
Db 2437 CATCTTGCAGACGGGCTCTTACTGTCCCGTGGCCTTCTCAGCTTCCGCTCCATGCT 2496
Qy 1392 GGGCTCTTCCCTGTACGCCCGGAGGCGTCAAGTCTGTCTCTGTGTGGTGTGCTGCCCT 1451
Db 2497 GGGCTCTTCCCTGTACGCCCGGAGGCGTCAAGTCTGTCTCTGTGTGGTGTGCTGCCCT 2556
Qy 1452 GCCTGCTGCTCAACCCACTGTGTACTGTCTTTCAACCCCACTTTCCGGGATGACCT 1511
Db 2557 GCCTGCTGCTCAACCCACTGTGTACTGTCTTTCAACCCCACTTTCCGGGATGACCT 2616
Qy 1512 TCGGGGCTTTCGGGCCCGCGCAGGGGACTCAGGGCCCTTAGCCTATGCTGCGGCGGGGA 1571
Db 2617 TCGGGGCTTTCGGGCCCGCGCAGGGGACTCAGGGCCCTTAGCCTATGCTGCGGCGGGGA 2676
Qy 1572 GCTGAGAGAGCTCTGTGATTCTACCCAGGCCCTGGTAGCCTTCTCTGATGTGGATCT 1631
Db 2677 GCTGAGAGAGCTCTGTGATTCTACCCAGGCCCTGGTAGCCTTCTCTGATGTGGATCT 2736
Qy 1632 CATCTGGAAGCTTCTGAAGCTGGGCGGCCCTTGGGCTGGAGACCTATGCTTCCCTC 1691
Db 2737 CATCTGGAAGCTTCTGAAGCTGGGCGGCCCTTGGGCTGGAGACCTATGCTTCCCTC 2796
Qy 1692 AGTGACCTCATCTCTGTCTAGCAGCCAGGGGCCCGCCAGGCTGGAGGCGACCATTTGT 1751
Db 2797 AGTGACCTCATCTCTGTCTAGCAGCCAGGGGCCCGCCAGGCTGGAGGCGACCATTTGT 2856
Qy 1752 AGAGCCAGAGGGGAACCACTTTGGGAAACCCCAACCTTCCATGATGAGAACTGCTGCT 1811
Db 2857 AGAGCCAGAGGGGAACCACTTTGGGAAACCCCAACCTTCCATGATGAGAACTGCTGCT 2916
Qy 1812 GAGGCGCAGAGGATCTACGCCAGCAGGTGGAGGCTTTGTACGGGGGTGGCGCTTTACGCC 1871
Db 2917 GAGGCGCAGAGGATCTACGCCAGCAGGTGGAGGCTTTGTACGGGGGTGGCGCTTTACGCC 2976
Qy 1872 CTCTGGCTTGGCCTTTGCTTACACAGTG 1899
Db 2977 CTCTGGCTTGGCCTTTGCTTACACAGTG 3004

Search completed: April 14, 2005, 19:13:39
Job time : 724.624 secs

This Page Blank (uspto)

GenCore version 5.1.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 14, 2005, 11:17:08 ; Search time 5432.4 Seconds
(without alignments)
16938.466 Million cell updates/sec

Title: US-10-664-667-6
Perfect score: 1899
Sequence: 1 aatagactactataggga.....tggccttgcttcacagtg 1899

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 4708233 seqs, 24227607955 residues

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : GenEmbl.*
1: gb_ba.*
2: gb_hcg.*
3: gb_in.*
4: gb_om.*
5: gb_ov.*
6: gb_pat.*
7: gb_ph.*
8: gb_pl.*
9: gb_pr.*
10: gb_ro.*
11: gb_ats.*
12: gb_sy.*
13: gb_un.*
14: gb_vl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1898	99.9	1899	6	AX301827
2	1898	99.9	2486	6	AX301825
3	1820.4	95.9	2322	6	AX098217
4	1811.6	95.4	3451	6	CQ840814
5	1754.4	92.4	2208	6	AX301830
6	1754.4	92.4	2711	6	AX301828
7	1754.4	92.4	2901	6	AX301833
8	1754.4	92.4	3325	9	AX358119
9	1754.4	92.4	3492	6	AX301831
10	1752.8	92.3	3119	6	AX549295
11	1752.8	92.3	3119	9	AB049405
12	1751.2	92.2	3429	6	AX451929
13	1749.6	92.1	2487	6	AX926297
14	1748	92.0	3273	6	CQ870634
15	1748	92.0	3286	9	BC047905
16	1748	92.0	3306	6	CQ840811
17	1746.4	92.0	3349	6	AX926300
18	1740.8	91.7	3041	6	E42917
19	1740.8	91.7	3041	6	AX016191

20	1740.8	91.7	3381	6	AX527920	Sequence
21	1738.2	91.5	2484	9	AF190501	Homo sapi
22	1365.4	71.9	1987	6	E42916	Novel gonad
23	1365.4	71.9	1987	6	AX016189	Sequence
24	1336	70.4	2901	6	AX301824	Sequence
25	1336	70.4	3637	6	AX301822	Sequence
26	1254.8	66.1	1272	9	AB083616	Homo sapi
27	1254.8	66.1	119596	6	AX926299	Sequence
28	1254.8	66.1	177344	9	AL356953	Human DNA
29	1189.6	62.6	1681	6	BD159457	Primer fo
30	1189.6	62.6	1681	6	AX882015	Sequence
31	1189.6	62.6	1681	6	BD012484	Guanosine
32	1189.6	62.6	1681	9	AC027377	Homo sapi
33	1186	62.5	178757	2	AC074007	Homo sapi
34	1051	55.3	2044	10	BC026896	Mus muscu
35	952	50.1	253599	2	AC117827	Mus muscu
36	915.6	48.2	227401	2	AC105613	Rattus no
37	915.6	48.2	238582	2	AC098661	Rattus no
38	915.6	48.2	247017	2	AC110384	Rattus no
39	723.2	38.1	1173	6	BD144308	Novel G-p
40	710.2	37.4	785	6	BD146757	Primer fo
41	710.2	37.4	785	6	AX866695	Sequence
42	513.4	27.0	515	6	E42913	Novel gonad
43	513.4	27.0	515	6	AX016183	Sequence
44	439.4	23.1	453	6	AR240484	Sequence
45	418.4	22.0	431	6	CQ724861	Sequence

ALIGNMENTS

RESULT 1
AX301827
LOCUS AX301827 1899 bp DNA linear PAT 30-NOV-2001
DEFINITION Sequence 6 from Patent WO0185768.
ACCESSION AX301827
VERSION AX301827.1 GI:17382885
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1
AUTHORS Gu, W.
TITLE Novel g-protein coupled receptors and uses therefor
JOURNAL Patent: WO 0185768-A 6 15-NOV-2001;
Millennium Pharmaceuticals, Inc. (US)
FEATURES
Location/Qualifiers
source
1..1899
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
<1..>1899
/note="unnamed protein product"
/codon_start=1
/protein_id="CAD13078.1"
/db_xref="GI:17382886"
/translation="NTTHYRESWVACVRSIGPGSTHASVRSOGLSLPAHPASIALAL
AASNTASGKLELDTFQSLQSLQALDLSWNAIRSIHFEARSTLSLVKLDITDQLTT
LPALGLGLMHLKLGKLNALQAFSKDSFKRLILEVFPYAYCCPYGNCAKSFQKASGQ
WEADLHLDDEESKRLGLARQENHYDQDLQLEMEBSKHPFSVQSGPTGPF
KPCYLPESGIRLAWAI VLLSVLCNGLVLLTVFAGGPAPLPVPKVFVGAAGANTL
TGISCGLLASVDALTFCQSEYGARWETGLGCRATGFLAVLGSEASVLLLTAAVQCS
VSVSCVAYGKSPSLGSRVAGLGLALAGLALPLASVGEYGCASPLCLPYAPEGQ
PALGFTVLVMMNSFCFLVAGAYIKLYCDLPRGDFEAVWDCWNRHRLWLIPIADGL
LYCPVAFSLFASMLGLFVTPPEAVKSVLLVLPACINFLYLLPLFRDRLRLR
LPCVDSGLAYAAAGELEKSCSTQALVAFSDVDLILEASEAGRPGLTYGPPSVT
LISCOQPGAPLEGSHCPEGNHFGNPQPSMDGELLRLAEGSTPAGGSLGGGQFP
SGLAFASHV"

ORIGIN

Query Match 99.9%; Score 1898; DB 6; Length 1899;
Best Local Similarity 100.0%; Pred. No. 0;

Qy	661	TGTGAGTACCTCTTTTGAAGACTGGGGCATCCGCGCTGGCGGTGTGGGCATCGTGTGCTC	720
Db	662	TGTGAGTACCTCTTTTGAAGACTGGGGCATCCGCGCTGGCGGTGTGGGCATCGTGTGCTC	721
Qy	721	TCCGTCGCTCTGCAATGGACTTGGTCTGCTGACCGTGTTCGTGGCGGGCTGCCCCCTG	780
Db	722	TCCGTCGCTCTGCAATGGACTTGGTCTGCTGACCGTGTTCGTGGCGGGCTGCCCCCTG	781
Qy	781	CCCCCGGTCAAGTTTGTGTGAGTGCAGTTGCAAGCGCCCAACACTTGTACTGCAATTTC	840
Db	782	CCCCCGGTCAAGTTTGTGTGAGTGCAGTTGCAAGCGCCCAACACTTGTACTGCAATTTC	841
Qy	841	TGTGGCCCTTCTAGCCTCAGTCAGTGCATGCCCTGACCTTTGGTCAGTCTCTGATGACGAGCC	900
Db	842	TGTGGCCCTTCTAGCCTCAGTCAGTGCATGCCCTGACCTTTGGTCAGTCTCTGATGACGAGCC	901
Qy	901	CGCTGGAGACGGGGCTTAGCTGTCGGGGCACCTGGCTTCTTGGCAGTACTTGGGTGGAG	960
Db	902	CGCTGGAGACGGGGCTTAGCTGTCGGGGCACCTGGCTTCTTGGCAGTACTTGGGTGGAG	961
Qy	961	GCATCGGTGCTGCTCTCACTCTGGCGCAGTGCAGTCAGCGTCTCCGTCTCCTGTGTC	1020
Db	962	GCATCGGTGCTGCTCTCACTCTGGCGCAGTGCAGTCAGCGTCTCCGTCTCCTGTGTC	1021
Qy	1021	CGGGCCTATGGGAAGTCCCTCCCTTGGGAGCGTTTCGAGACGGGTCTTAGGCTGCCTG	1080
Db	1022	CGGGCCTATGGGAAGTCCCTCCCTTGGGAGCGTTTCGAGACGGGTCTTAGGCTGCCTG	1081
Qy	1081	GCACGGCAGGGCTGGCGCGACCTGCCCTCGGCTCAGTGGGAGAAACAGGGGCTCC	1140
Db	1082	GCACGGCAGGGCTGGCGCGACCTGCCCTCGGCTCAGTGGGAGAAACAGGGGCTCC	1141
Qy	1141	CCACTCTGCTGCCCTACCGCGCACCTGAGGGTCAGCCAGCAGCCCTTGGGCTTCAACGGTG	1200
Db	1142	CCACTCTGCTGCCCTACCGCGCACCTGAGGGTCAGCCAGCAGCCCTTGGGCTTCAACGGTG	1201
Qy	1201	GCCTCGGTGANTGAATCTCTTCTGTTTCTGGTCTGGCGGGTGCCTACATCAAACTG	1260
Db	1202	GCCTCGGTGANTGAATCTCTTCTGTTTCTGGTCTGGCGGGTGCCTACATCAAACTG	1261
Qy	1261	TACTGTGACTGCGCGGGGGGACCTTTGAGGCGCTGTGGACCTGCCCATGGTGAGGCAC	1320
Db	1262	TACTGTGACTGCGCGGGGGGACCTTTGAGGCGCTGTGGACCTGCCCATGGTGAGGCAC	1321
Qy	1321	GTGGCTTGGCTCATCTTTCGACAGCGGGCTCCTCTACTGTCGGCTGGCTTCTCAGCTTC	1380
Db	1322	GTGGCTTGGCTCATCTTTCGACAGCGGGCTCCTCTACTGTCGGCTGGCTTCTCAGCTTC	1381
Qy	1381	GCCTCCATGTGGGCTCTTCCCTGTGTCACGCCGAGGCGGTCAAGTCTGCTGCTGGTG	1440
Db	1382	GCCTCCATGTGGGCTCTTCCCTGTGTCACGCCGAGGCGGTCAAGTCTGCTGCTGGTG	1441
Qy	1441	GTGCTGCCCCCTGCTGCTGCTCAACCCACTGCTGTACTGTCTTTCACCCCCCACTTC	1500
Db	1442	GTGCTGCCCCCTGCTGCTGCTCAACCCACTGCTGTACTGTCTTTCACCCCCCACTTC	1501
Qy	1501	CGGGATGACTCTTCGGCGGCTTCGGCCCCCGCAGAGGGACTCAGGGGCCCTTAGCCTATGCT	1560
Db	1502	CGGGATGACTCTTCGGCGGCTTCGGCCCCCGCAGAGGGACTCAGGGGCCCTTAGCCTATGCT	1561
Qy	1561	CGGGCCGGGAGCTGGAGAAGAGCTCTGTGATTTCTACCCAGGCCCTTGGTTCCTCTCT	1620
Db	1562	CGGGCCGGGAGCTGGAGAAGAGCTCTGTGATTTCTACCCAGGCCCTTGGTTCCTCTCT	1621
Qy	1621	GATGTGGATCTCATTTCTGGAAGCTTCTGAAGCTGGCGGGCCCCCTGGGCTGGAGACTAT	1680
Db	1622	GATGTGGATCTCATTTCTGGAAGCTTCTGAAGCTGGCGGGCCCCCTGGGCTGGAGACTAT	1681
Qy	1681	GGCTTCCCCTCAGTGACCTCATCTCTGTCAGACCCAGGGGCCCCCCAGGCTGGAGGGC	1740
Db	1682	GGCTTCCCCTCAGTGACCTCATCTCTGTCAGACCCAGGGGCCCCCCAGGCTGGAGGGC	1741
Qy	1741	AGCCATTGTGTAGAGCCAGAGGGGAACACTTTTGGGAAACCCCCAAACCTCTCATGTGATGA	1800

Db 1742 AGCCATTGTGTAGAGCCAGAGGGAAACACTTTTGGGAAACCCCAACCCCTCCATGGATGGA 1801
Qy 1801 GAACCTGCTGTAGAGGCGAGAGGACTAGCCACAGAGTGGAGGCTTCTCAGGGGGTGGC 1860
Db 1802 GAACCTGCTGTAGAGGCGAGAGGACTAGCCACAGAGTGGAGGCTTCTCAGGGGGTGGC 1861
Qy 1861 GGCTTTACAGCCCTCTGGCTTGGCCCTTGTCTTACACAGTG 1899
Db 1862 GGCTTTACAGCCCTCTGGCTTGGCCCTTGTCTTACACAGTG 1900

RESULT 3
AX098217
LOCUS AX098217 2322 bp DNA linear PAT 30-MAR-2001
DEFINITION Sequence 129 from Patent WO0118542.
ACCESSION AX098217
VERSION AX098217.1 GI:13515330
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1
Lee, J., Thompson, P. and Lillie, J.
IDENTIFICATION, assessment, prevention, and therapy of ovarian
cancer
JOURNAL Patent: WO 0118542-A 129 15-MAR-2001;
AUTHORS Millennium Predictive Medicine, Inc. (US)
TITLE Location/Qualifiers
FEATURES
source 1..2322
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

ORIGIN
Query Match 95.9%; Score 1820.4; DB 6; Length 2322;
Best Local Similarity 99.9%; Pred No. 0;
Matches 1832; Conservative 0; Mismatches 1; Indels 1; Gaps 1;
Qy 66 CCACGCGTCCGTGGAGCGAGCCAGGGTCTGAGCCTGCCGGCTCATCCAGCCTCTCTTGC 125
Db 1 CCACGCGTCCGTGGAGCGAGCCAGGGTCTGAGCCTGCCGGCTCATCCAGCCTCTCTTGC 60
Qy 126 TGCCCTACGCGCTCCACACACACACCGCATCTGGGAAATTGGAGTNGACACCTTCAGCCA 185
Db 61 TGCCCTACGCGCTCCACACACACACCGCATCTGGGAAATTGGAGCT-GACACCTTCAGCCA 119
Qy 186 GCTGAGCTCCCTGCAAGCCCTGGATCTTTAGCTGGAAAGCCATCGGTCATCCACCTTGA 245
Db 120 GCTGAGCTCCCTGCAAGCCCTGGATCTTTAGCTGGAAAGCCATCGGTCATCCACCTTGA 179
Qy 246 GGCTTTCTCCACCTTGCACTCCCTGGTCAAGCTGGACCTGACAGAACACCAAGCTGACCC 305
Db 180 GGCTTTCTCCACCTTGCACTCCCTGGTCAAGCTGGACCTGACAGAACACCAAGCTGACCC 239
Qy 306 ACTGCCCTCGGTGACCTTTGGGGGCTTGAATGATCTGAAGCTCAAGAGAACTTGTCTCT 365
Db 240 ACTGCCCTCGGTGACCTTTGGGGGCTTGAATGATCTGAAGCTCAAGAGAACTTGTCTCT 299
Qy 366 CTCCAGGCTTCTCCAGGACAGTTTCCAAAACCTGAGGATCTCTGGAGTGCCTTATGC 425
Db 300 CTCCAGGCTTCTCCAGGACAGTTTCCAAAACCTGAGGATCTCTGAAGTGCCTTATGC 359
Qy 426 CTACAGTGTCTCTCCATATGGATGTGTGCCAGCTTCTTCAAGGCTCTGGGCGAGTGGGA 485
Db 360 CTACAGTGTCTCTCCATATGGATGTGTGCCAGCTTCTTCAAGGCTCTGGGCGAGTGGGA 419
Qy 486 GGCTGAAGACCTTCACTTGTATGATGAGGAGTCTTCAAAAAGGCCCTTGGGCTCTCTTGC 545
Db 420 GGCTGAAGACCTTCACTTGTATGATGAGGAGTCTTCAAAAAGGCCCTTGGGCTCTCTTGC 479
Qy 546 CAGACAAGCAGAACCACTATGACCAGGACCTTGGATGAGCTCCAGCTGGAGATGGAGGA 605

Db 480 CAGACAAGCAGAACCACTATGACCAGGACCTTGGATGAGCTCCAGCTGGAGATGGAGGA 539
Qy 606 CTCAAAAGCCACACCCAGTGTCCAGTGTAGCCCTACTCCAGGCCCTTCAAGGCCCTGTGA 665
Db 540 CTCAAAAGCCACACCCAGTGTCCAGTGTAGCCCTACTCCAGGCCCTTCAAGGCCCTGTGA 599
Qy 666 GTACCTCTTTGAAAGCTGGGGCATCCGCTGGCCGCTGTGGGCCATCGTGTGTCTCTCGT 725
Db 600 GTACCTCTTTGAAAGCTGGGGCATCCGCTGGCCGCTGTGGGCCATCGTGTGTCTCTCGT 659
Qy 726 GCTCTGCAATGGACTGTGTGTGTGACCGTGTTCGCTGGGGGGCTGCCCCCTTGCCTCC 785
Db 660 GCTCTGCAATGGACTGTGTGTGTGACCGTGTTCGCTGGGGGGCTGCCCCCTTGCCTCC 719
Qy 786 GGTCAGCTTTGT 845
Db 720 GGTCAGCTTTGT 779
Qy 846 CCTTCTAGCCTCAGTCGATGCCCTGACCTTTGGTTCAGTTCCTCTGAGTACGAGAGCCCGCTG 905
Db 780 CCTTCTAGCCTCAGTCGATGCCCTGACCTTTGGTTCAGTTCCTCTGAGTACGAGAGCCCGCTG 839
Qy 906 GGAGACGGGGCTAGGGCTGCGGGCCACTGGCTTCTCTGGCAGTACTTGGGTCTGGAGGCATC 965
Db 840 GGAGACGGGGCTAGGGCTGCGGGCCACTGGCTTCTCTGGCAGTACTTGGGTCTGGAGGCATC 899
Qy 966 GGTGCTGCTGTCTACTCTGGCCGAGTGCAGTGTGAGCGTCTCTCGTCTCTGTGTCCGGGC 1025
Db 900 GGTGCTGCTGTCTACTCTGGCCGAGTGCAGTGTGAGCGTCTCTCGTCTCTGTGTCCGGGC 959
Qy 1026 CTATGGGAAGTCCCTCCCTGGGACGCTTCGAGCAGGGGTCTTAGGCTGCCTGGGCACT 1085
Db 960 CTATGGGAAGTCCCTCCCTGGGACGCTTCGAGCAGGGGTCTTAGGCTGCCTGGGCACT 1019
Qy 1086 GGAGGGCTGCGCGCCGCACTTGCCCTGGCTCAGTGGGAGAAATACGGGGCTCTCCCACT 1145
Db 1020 GGAGGGCTGCGCGCCGCACTTGCCCTGGCTCAGTGGGAGAAATACGGGGCTCTCCCACT 1079
Qy 1146 CTGCTGCTGCTAGCGGCCACTGAGGGTTCAGCAGAGCCCTGGGGTTCACCGTGGGCT 1205
Db 1080 CTGCTGCTGCTAGCGGCCACTGAGGGTTCAGCAGAGCCCTGGGGTTCACCGTGGGCT 1139
Qy 1206 GGTGATGATCAATCCCTTCTTCTGTTTCCGTGCTGGCCGCTGCTACATCAAACTGACTG 1265
Db 1140 GGTGATGATCAATCCCTTCTTCTGTTTCCGTGCTGGCCGCTGCTACATCAAACTGACTG 1199
Qy 1266 TGACCTGCGCGGGGCACTTTTGAAGCCGTGTGGGACTGGGCCATGCTGAGGCACGCTGGC 1325
Db 1200 TGACCTGCGCGGGGCACTTTTGAAGCCGTGTGGGACTGGGCCATGCTGAGGCACGCTGGC 1259
Qy 1326 CTGGCTCATTTTGCAGACGGGCTCTCTACTGTTCGCTGGCTTCTCTAGCTTTCGGCTC 1385
Db 1260 CTGGCTCATTTTGCAGACGGGCTCTCTACTGTTCGCTGGCTTCTCTAGCTTTCGGCTC 1319
Qy 1386 CATGCTGGGCTCTTTCCTGTACGCCCGAGGGCGCTCAAGTCTGTCTGTGTGTGTGTGTGT 1445
Db 1320 CATGCTGGGCTCTTTCCTGTACGCCCGAGGGCGCTCAAGTCTGTCTGTGTGTGTGTGTGT 1379
Qy 1446 GCGCTGCTGCTGCTCAACCCACTGCTGTACTCTCTTCAAAACCCCACTTCCCGGA 1505
Db 1380 GCGCTGCTGCTGCTCAACCCACTGCTGTACTCTCTTCAAAACCCCACTTCCCGGA 1439
Qy 1506 TGACCTTTCGGCGGCTTCGGCCCGCGCAGGGGACTCAGGGCCCTTAGCCTATGCTCGGC 1565
Db 1440 TGACCTTTCGGCGGCTTCGGCCCGCGCAGGGGACTCAGGGCCCTTAGCCTATGCTCGGC 1499
Qy 1566 CGGGGAGCTGGAGAGCTCTCTGTGATTTCTACCCAGGCCCTTGTAGCCTTCTGTATGT 1625
Db 1500 CGGGGAGCTGGAGAGCTCTCTGTGATTTCTACCCAGGCCCTTGTAGCCTTCTGTATGT 1559
Qy 1626 GGATCTCATTTCTGAAGCTTCTGAAGCTGGGGCGGCCCTTGGGCTGGAGACCTATGCTT 1685

Db 1560 GGATCTCATCTTGGAGCTTCTGAAGCTGGGGGGCCCCCTGGGCTGGAGACCTATGGCTT 1619

Qy 1686 CCCCTCAGTGACCTCATCTCTGTGTCAGCAGCAGGGGGCCCCCAGGCTGGAGGAGCCCA 1745

Db 1620 CCCCTCAGTGACCTCATCTCTGTGTCAGCAGCAGGGGGCCCCCAGGCTGGAGGAGCCCA 1679

Qy 1746 TTGTGTAGAGCAGAGGGGAAACCACTTTGGGAAACCCCAACCTCGATGGATGGAGAACT 1805

Db 1680 TTGTGTAGAGCAGAGGGGAAACCACTTTGGGAAACCCCAACCTCGATGGATGGAGAACT 1739

Qy 1806 GCTGCTAGAGCAGAGGGGATCTACGCCAGCAGCTGGAGCTTGTTCAGGGGGTGGCGCTT 1865

Db 1740 GCTGCTAGAGCAGAGGGGATCTACGCCAGCAGCTGGAGCTTGTTCAGGGGGTGGCGCTT 1799

Qy 1866 TCAGCCCTCTGCTTGGCTTGGCTTGCCTTACACAGTG 1899

Db 1800 TCAGCCCTCTGCTTGGCTTGGCTTGCCTTACACAGTG 1833

RESULT 4

LOCUS CQ840814 3451 bp DNA linear PAT 29-JUL-2004

DEFINITION Sequence 4 from Patent WO2004058818.

ACCESSION CQ840814

VERSION CQ840814.1 GI:50838404

KEYWORDS

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1

REFERENCE

AUTHORS Muda, M.

TITLE variants of lgr6

JOURNAL Patent: WO 2004058818-A 4 15-JUL-2004;

Applied Research Systems ABS Holding N.V. (AN)

FEATURES

Location/Qualifiers

1..3451

/organism="Homo sapiens"

/mol_type="unassigned DNA"

/db_xref="taxon:9606"

47..1093

/note="unnamed protein product"

/codon_start=1

/protein_id="CAH05694.1"

/db_xref="GI:50838405"

/translation="MGRPLTLVQVSIISARDLSMNLTLOPLPHLRLBELR

LSGNHLSHPGQAFSLYSLKTLMLQNNQGGIPAEALWELPSLOSRLDANLISLP

ERSFGLSLRLHLDDNALTEIPVRALNNLPALQAMTLNRIHIEDYAFQNLTSL

VVLHNNRKHQHGTHSEGLHNLETLDLNLYNKLOEPFVAINTLRLQELGPHNNIK

AIPEKAFNGLPQTHLYDNPQFVRSFAQYLPKLHLSLNGAMDIOEPFDLKGTT

SLIEILTRAGIRLLPSGMQQLPRLRLVLELSHNQIEBLSLHRCQKLEMRSGGLYF

LPPSARGAWSGARV"

ORIGIN

Query Match 95.4%; Score 1811.6; DB 6; Length 3451;

Best Local Similarity 99.7%; Pred. No. 0;

Matches 1825; Conservative 0; Mismatches 4; Indels 1; Gaps 1;

Qy 70 GGTCCGTGGAGCGGAGCCAGGCTGAGCCTGCCGGCTCATCCAGCCTCTTCTGCTGC 129

Db 1066 GGGGGCATTGAGCGGAGCGAGGCTGAGCCTGCCGGCTCATCCAGCCTCTTCTGCTGC 1125

Qy 130 CTAGCGGCTCCAAACACACCCGATCTGGGAAATTGGAGCTNGACACTTCAGCCAGCTG 189

Db 1126 CTAGCGGCTCCAAACACACCCGATCTGGGAAATTGGAGCT-GACACTTCAGCCAGCTG 1184

Qy 190 AGCTCCCTGCAAGCCCTTGGATCTTAGCTGGAAACCGCATCCGGTCCATCCACCTCAGGCC 249

Db 1185 AGCTCCCTGCAAGCCCTTGGATCTTAGCTGGAAACCGCATCCGGTCCATCCACCTCAGGCC 1244

Qy 250 TTCTCAGCCCTGATCCCTGCTCAAGCTGGAGCTGACAGAACACAGCTGACCACTG 309

Db 1245 TTCTCAGCCCTGATCCCTGCTCAAGCTGGAGCTGACAGAACACAGCTGACCACTG 1304

Qy 310 CCCCTGGCTGACCTTGGGGGCTTGATGCATCTGAAGCTCAAAGGGAACCTTGTCTCTCTCC 369

Db 1305 CCCCTGGCTGACCTTGGGGGCTTGATGCATCTGAAGCTCAAAGGGAACCTTGTCTCTCTCC 1364

Qy 370 CAGGCTTCTCCNAGGACAGTTTCCCAAACTGAGGATCCTGGAGGTGCCCTATGCGCTAC 429

Db 1365 CAGGCTTCTCCNAGGACAGTTTCCCAAACTGAGGATCCTGGAGGTGCCCTATGCGCTAC 1424

Qy 430 CAGTCTCTCCCTATGGGATGTGTGCCAGCTTCTTCAAGGGCTCTTGGGACAGTGGAGGCT 489

Db 1425 CAGTCTCTCCCTATGGGATGTGTGCCAGCTTCTTCAAGGGCTCTTGGGACAGTGGAGGCT 1484

Qy 490 GAAGACCTTCACTTGTATGATGAGAGGCTTCAAAGGCGCCCTGGGCTCTCTTCCAGAGA 549

Db 1485 GAAGACCTTCACTTGTATGATGAGAGGCTTCAAAGGCGCCCTGGGCTCTCTTCCAGAGA 1544

Qy 550 CAAGCAGAGAACCACTATGACAGGACCTGATGAGCTTCCAGCTGGAGATGGAGGACTCA 609

Db 1545 CAAGCAGAGAACCACTATGACAGGACCTGATGAGCTTCCAGCTGGAGATGGAGGACTCA 1604

Qy 610 AAGCCACACCCAGTGTCCAGTGTAGCCCTTACTCCAGGCCCCCTTCAAGCCCCCTGTGAGTAC 669

Db 1605 AAGCCACACCCAGTGTCCAGTGTAGCCCTTACTCCAGGCCCCCTTCAAGCCCCCTGTGAGTAC 1664

Qy 670 CTCTTTGAAAGCTGGGGCATCCGCTGCGCTGTGGGCCATCGTGTCTCTCTGCTGCTC 729

Db 1665 CTCTTTGAAAGCTGGGGCATCCGCTGCGCTGTGGGCCATCGTGTCTCTCTGCTGCTC 1724

Qy 730 TGCAATGACAGTGTGCTGCTGACCGTGTTCGCTGGCGGGCTGCCCCCTGCCCCGGGTC 789

Db 1725 TGCAATGACAGTGTGCTGCTGACCGTGTTCGCTGGCGGGCTGCCCCCTGCCCCGGGTC 1784

Qy 790 AAGTTTGTGTAGTGTGCGATTGAGGCGCCCAACACTTGACTGGCATTTCTGTGGCCTT 849

Db 1785 AAGTTTGTGTAGTGTGCGATTGAGGCGCCCAACACTTGACTGGCATTTCTGTGGCCTT 1844

Qy 850 CTAGCCTCAGTGTGATGCTGACCTTTGGTTCAGTTCCTGAGTACGAGGCGGCTGGAG 909

Db 1845 CTAGCCTCAGTGTGATGCTGACCTTTGGTTCAGTTCCTGAGTACGAGGCGGCTGGAG 1904

Qy 910 ACGGGCTAGGCTGCGGGGCCACTGGCTTCTGGCAGTACTTGGGTGGAGGACATCGGTG 969

Db 1905 ACGGGCTAGGCTGCGGGGCCACTGGCTTCTGGCAGTACTTGGGTGGAGGACATCGGTG 1964

Qy 970 CTGCTGCTCACTCTGGCGCGCAGTGCAGTGCAGCGTCTCCGTCTCTGTGTCGGGCTAT 1029

Db 1965 CTGCTGCTCACTCTGGCGCGCAGTGCAGTGCAGCGTCTCCGTCTCTGTGTCGGGCTAT 2024

Qy 1030 GGGAGTCCCCCTCCCTGGGAGCGTTCCAGCAGGGGCTCTAGGCTGCTGGCACTGGCA 1089

Db 2025 GGGAGTCCCCCTCCCTGGGAGCGTTCCAGCAGGGGCTCTAGGCTGCTGGCACTGGCA 2084

Qy 1090 GGGCTGGCGCGGACACTGCCCCCTGGCTTCTGGCAGTACTTGGGTGGAGGACATCGGTG 1149

Db 2085 GGGCTGGCGCGGACACTGCCCCCTGGCTTCTGGCAGTACTTGGGTGGAGGACATCGGTG 2144

Qy 1150 CTGCCCTACGCGCCACTGAGGGTTCAGCAGCGGCTTGGGCTTCAACGTCGGGCTGGTG 1209

Db 2145 CTGCCCTACGCGCCACTGAGGGTTCAGCAGCGGCTTGGGCTTCAACGTCGGGCTGGTG 2204

Qy 1210 ATGATGAACCTCTTCTGTTCTGCTGGTGGCGGCTTACATCAAACTGTACTGTGAC 1269

Db 2205 ATGATGAACCTCTTCTGTTCTGCTGGTGGCGGCTTACATCAAACTGTACTGTGAC 2264

Qy 1270 CTGCCCGGGGGGAGCTTTGAGCGGCTGTGGGACATGGGCGCATGGTGGAGCAGTGGCCTGG 1329

Db 2265 CTGCCCGGGGGGAGCTTTGAGCGGCTGTGGGACATGGGCGCATGGTGGAGCAGTGGCCTGG 2324

Qy 1330 CTCACTTCGAGAGGGGCTCTCTACTGTCGGTGGGCTTCCCTCAGCTTCGCTCCATG 1389

Db 2325 CTCACTTCGAGAGGGGCTCTCTACTGTCGGTGGGCTTCCCTCAGCTTCGCTCCATG 2384

QY	1092	GCTGGCGCGGCACTGCCCTCGCTCAGTGGGAGAAATACGGGGCTCCCACTCTGCCT	1151
Db	1401	GCTGGCGCGGCACTGCCCTCGCTCAGTGGGAGAAATACGGGGCTCCCACTCTGCCT	1460
QY	1152	GCCTACGCGCGCACTGAGGGTCAGCAGACGCGCTTGCAGTGGGCGGCTGGTAT	1211
Db	1461	GCCTACGCGCGCACTGAGGGTCAGCAGACGCGCTTGCAGTGGGCGGCTGGTAT	1520
QY	1212	GATGAATCTCTTCTGTTTCTGTTCTGTTGCGGCGGTGCTACATCAAACTGTACTGTGACCT	1271
Db	1521	GATGAATCTCTTCTGTTTCTGTTTCTGTTGCGGCGGTGCTACATCAAACTGTACTGTGACCT	1580
QY	1272	GCGCGGGGCGACTTTGAGGCGGTGTGGACCTCGGCATGTGTGAGCAGCTGGCTGGCT	1331
Db	1581	GCGCGGGGCGACTTTGAGGCGGTGTGGACCTCGGCATGTGTGAGCAGCTGGCTGGCT	1640
QY	1332	CATCTTTCGAGACGGGCTCTCTACTGTCTGCTGCGTGGCTTCTTCACGTTCCGCTCCATGCT	1391
Db	1641	CATCTTTCGAGACGGGCTCTCTACTGTCTGCTGCGTGGCTTCTTCACGTTCCGCTCCATGCT	1700
QY	1392	GCGCTCTTCTCTGTACGCGCGAGGCGGTCAAGTCTGTCTGCTGGTGGTCTGCCCT	1451
Db	1701	GCGCTCTTCTCTGTACGCGCGAGGCGGTCAAGTCTGTCTGCTGGTGGTCTGCCCT	1760
QY	1452	GCTGCTGCTCAACCCACTGCTGTACTGCTCTTCAACCCCACTTCCGGGATGACCT	1511
Db	1761	GCTGCTGCTCAACCCACTGCTGTACTGCTCTTCAACCCCACTTCCGGGATGACCT	1820
QY	1512	TGGCGGGCTTCGGCGCGGCGAGGACTCAGGGCGGCTAGGCTATGCTGCGGCGGGGA	1571
Db	1821	TGGCGGGCTTCGGCGCGGCGAGGACTCAGGGCGGCTAGGCTATGCTGCGGCGGGGA	1880
QY	1572	GCTGAGAGAGCTCTGTGATTTACCCAGGCGGCTGGTAGCTTCTCTGATGTGGATCT	1631
Db	1881	GCTGAGAGAGCTCTGTGATTTACCCAGGCGGCTGGTAGCTTCTCTGATGTGGATCT	1940
QY	1632	CATTCTGAAGCTTCTGAAGCTGGGCGGCGGCTGGTAGGCTATGAGCTTCCCTC	1691
Db	1941	CATTCTGAAGCTTCTGAAGCTGGGCGGCGGCTGGTAGGCTATGAGCTTCCCTC	2000
QY	1692	AGTGACCTCATCTCTGTACAGCAGGAGGCGGCGGCTGGTAGGCTATGAGCTTCCCTC	1751
Db	2001	AGTGACCTCATCTCTGTGTACAGCAGGAGGCGGCGGCTGGTAGGCTATGAGCTTCCCTC	2060
QY	1752	AGAGCCAGAGGGGAACCACTTTGGGAACCCCAACCTCCATGGATGGAGAACTCTGCT	1811
Db	2061	AGAGCCAGAGGGGAACCACTTTGGGAACCCCAACCTCCATGGATGGAGAACTCTGCT	2120
QY	1812	GAGGCGAGAGGATCTACGCCAGAGGTGGAGGCTTGTACAGGGGTGGCGGCTTTCAGCC	1871
Db	2121	GAGGCGAGAGGATCTACGCCAGAGGTGGAGGCTTGTACAGGGGTGGCGGCTTTCAGCC	2180
QY	1872	CTCTGGCTTGGCTTGGCTTTCACAGTG	1899
Db	2181	CTCTGGCTTGGCTTGGCTTTCACAGTG	2208

RESULT 6
AX301828
LOCUS AX301828 2711 bp DNA linear PAT 30-NOV-2001
DEFINITION Sequence 7 from Patent WO0185768.
ACCESSION AX301828
VERSION AX301828.1 GI:17382887
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS Gu, W.
TITLE Novel g-protein coupled receptors and uses therefor
JOURNAL Patent: WO 0185768-A 7 15-NOV-2001;
Millennium Pharmaceuticals, Inc. (US)

FEATURES
source

CDS

Location/Qualifiers
1..2711
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
41..2211
/note="unnamed protein product"
/codon_start=1
/protein_id="CAD13079.1"
/db_xref="GI:1738288"
/translation="GLHNLETLDLNLYNKQLQFPVAIRTLRQLBELGFHNNIKAIPEK
AFMGNPLQTHFYDNPQFVGRSAFYLPKHLNLSLNGAMDIDQFPDLKRTTSEIL
TLTRAGIRLLPSMGCCQQLPRURVLELSHNSHLELPSLHRCQKLEIGLQHNRIWELI
DTFQSLQALDLSWNAIRSHPEAFSTLSLVKLDLTQNLTLPLAGLGLMHLK
LKGMLASQAFSKDSFPKLRILEYPAVQCCPYGMCASFRRKASGOWEADLHLDDEES
SKRLPLARQENHVDQDLDELOLEMEDSKPHSVQCSPTGPKPKCEYLFPESWGIR
LAVNAIILLVNLGVLLTVTPAGGPAPLPVKVWVGAIGANTLTGTISGLLASVDA
LTFQFSEYGARWETIGCRATGFLAVLGEASVLLLLTAAVQSVSVSVRAYGKSP
SLGSVRAGVLGCLALAGLAALPLASVEYGAISPLCLPYAPPGGPALGFTVALVMM
NSFCFLVVAGAYIKLYCDLPRGDEAVMDCAMVRHVAWLIFADGLLYCPVAPLSPASM
LGLFPVTPPEAVKSVLLVPLPACLNPLLYLLFNPHFRDRLRLRPRAGDSGLPAA
AGELEKSSCDSTQALVAFSDVDLILEASEAGRPGLETYGPSPVTLISCOQPGAPRLE
GSHCVPEGNHFGNPQPSMDGELLRLRAGSIPAGGSLSGGGPQPSGLAPASHV"

ORIGIN

	Query Match	92.4%;	Score 1754.4;	DB 6;	Length 2711;
	Best Local Similarity	99.9%;	Pred. No. 0;		
	Matches 1766;	Conservative	0;	Mismatches	1; Indels 1; Gaps 1;
QY	132	AGCGGCTCCAAACAACCGCATCTGGGAAATTTGAGCTNGACACTTTCAGCAGCTGAG	191		
Db	442	ATCGGCTTCCAAACAACCGCATCTGGGAAATTTGAGCT-GACACCTTTCAGCAGCTGAG	500		
QY	192	CTCCCTGCAAGCCCTGGATCTTAGCTGGAACGCGCATCCGCTCCATCCACCTGAGGCTT	251		
Db	501	CTCCCTGCAAGCCCTGGATCTTAGCTGGAACGCGCATCCGCTCCATCCACCTGAGGCTT	560		
QY	252	CTCCACCTGCACCTCCCTGCTCAAGCTGGACCTTGACAGACAACAGCTGACACACTGCC	311		
Db	561	CTCCACCTGCACCTCCCTGCTCAAGCTGGACCTTGACAGACAACAGCTGACACACTGCC	620		
QY	312	CCTGCTGGAATTTGGGGCTTGTGATGATCTGAAGCTCAAGAGGAACCTTGTCTCTCCCA	371		
Db	621	CCTGCTGGAATTTGGGGCTTGTGATGATCTGAAGCTCAAGAGGAACCTTGTCTCTCCCA	680		
QY	372	GGCTTCTCCAGAGACAGTTTCCCAAACTCAGAGATCTCGAGGTGCTTATGCTATCA	431		
Db	681	GGCTTCTCCAGAGACAGTTTCCCAAACTCAGAGATCTCGAGGTGCTTATGCTATCA	740		
QY	432	GTGCTGTCCCTATGGGATGTGTGCCAGCTTCTTCAAGGCTCTTGGGAGTGGAGGCTGA	491		
Db	741	GTGCTGTCCCTATGGGATGTGTGCCAGCTTCTTCAAGGCTCTTGGGAGTGGAGGCTGA	800		
QY	492	AGACCTTCACTTGTATGATGAGGATCTTCAAAAGGCGCTTGGGCTCTCTTGCAGACA	551		
Db	801	AGACCTTCACTTGTATGATGAGGATCTTCAAAAGGCGCTTGGGCTCTCTTGCAGACA	860		
QY	552	AGCAGAGAACCACTATGACACGAGGACCTGGATGAGCTCCAGCTGGAGTGGAGGACTCAA	611		
Db	861	AGCAGAGAACCACTATGACACGAGGACCTGGATGAGCTCCAGCTGGAGTGGAGGACTCAA	920		
QY	612	GCCACACCCAGTGTCCAGTGTAGCCCTTACTCCAGGCGGCTTCAAGCCCTGTGAGTACT	671		
Db	921	GCCACACCCAGTGTCCAGTGTAGCCCTTACTCCAGGCGGCTTCAAGCCCTGTGAGTACT	980		
QY	672	CTTTGAAAGCTGGGCAATCCGCTTGGCGGTGTGGCCATCGTGTGCTCTCCGTGCTCTG	731		
Db	981	CTTTGAAAGCTGGGCAATCCGCTTGGCGGTGTGGCCATCGTGTGCTCTCCGTGCTCTG	1040		
QY	732	CAATGGACTGTGTGCTGACCGTGTTCGCTGGGCGGCTCCGCTTGGGCTTCAAGGCTCAA	791		
Db	1041	CAATGGACTGTGTGCTGACCGTGTTCGCTGGGCGGCTCCGCTTGGGCTTCAAGGCTCAA	1100		

QY 792 GTTGTGGTAGTGCGATTGACGGCGCCAAACACCTTGACTGCAATTTCTGTGGCTTCT 851
 Db 1101 GTTGTGGTAGTGCGATTGACGGCGCCAAACACCTTGACTGCAATTTCTGTGGCTTCT 1160
 QY 852 AGCCTCAGTCGATGCGCTGACCTTTGGTTCAGTTCTCTGAGTACGAGCCCGCTGGAGAC 911
 Db 1161 AGCCTCAGTCGATGCGCTGACCTTTGGTTCAGTTCTCTGAGTACGAGCCCGCTGGAGAC 1220
 QY 912 GGGGCTAGGCTGCGGGCCACTGGCTTCTGCGACTACTTGGTGGAGGACATCGGTGCT 971
 Db 1221 GGGGCTAGGCTGCGGGCCACTGGCTTCTGCGACTACTTGGTGGAGGACATCGGTGCT 1280
 QY 972 GCTGCTCACTCTGGCGCGAGTCAGTCAGCGCTCTCCGTCTCTGTGTCGGGCGCTATGG 1031
 Db 1281 GCTGCTCACTCTGGCGCGAGTCAGTCAGCGCTCTCCGTCTCTGTGTCGGGCGCTATGG 1340
 QY 1032 GAAGTCCCTTCCCTGGGACGCTTCGAGCAGGGGTCTTAGGCTGCCCTGGCACTGGCAGG 1091
 Db 1341 GAAGTCCCTTCCCTGGGACGCTTCGAGCAGGGGTCTTAGGCTGCCCTGGCACTGGCAGG 1400
 QY 1092 GCTGCGCGCGACTGCGCCCTGGCTCAGTGGGAGAAACGGGGCTCCCACTCTGCT 1151
 Db 1401 GCTGCGCGCGACTGCGCCCTGGCTCAGTGGGAGAAACGGGGCTCCCACTCTGCT 1460
 QY 1152 GCGCTACCGCGCACCTGAGGGTCAGCAGCAGCCCTGGGCTTCAACGCTGGCCCTGGTAT 1211
 Db 1461 GCGCTACCGCGCACCTGAGGGTCAGCAGCAGCCCTGGGCTTCAACGCTGGCCCTGGTAT 1520
 QY 1212 GATGAACCTCTTCTGTTTCTGCTGCTGGCGGCTTCTACATCAAACTGTACTGTGACCT 1271
 Db 1521 GATGAACCTCTTCTGTTTCTGCTGCTGGCGGCTTCTACATCAAACTGTACTGTGACCT 1580
 QY 1272 GCGCGGGCGACTTTGAGGCGGCTGTGGACTGCGGCTAGTGGGACGCTGGCGCTGGCT 1331
 Db 1581 GCGCGGGCGACTTTGAGGCGGCTGTGGACTGCGGCTAGTGGGACGCTGGCGCTGGCT 1640
 QY 1332 CATCTTCGACAGGGGCTCTCTACTGCTGCGTGGGCTTCTCAGCTTCGCTCCATGCT 1391
 Db 1641 CATCTTCGACAGGGGCTCTCTACTGCTGCGTGGGCTTCTCAGCTTCGCTCCATGCT 1700
 QY 1392 GGGGCTCTTCCCTGTACGCGCGAGGCGCTCAAGTCTGCTGCTGCTGGTGGTGGCTGCCCT 1451
 Db 1701 GGGGCTCTTCCCTGTACGCGCGAGGCGCTCAAGTCTGCTGCTGCTGGTGGTGGCTGCCCT 1760
 QY 1452 GCGTGCCTGCTCAACCCACTGCTGCTACTGCTTCTCAACCCCACTTCCGGGATGACCT 1511
 Db 1761 GCGTGCCTGCTCAACCCACTGCTGCTACTGCTTCTCAACCCCACTTCCGGGATGACCT 1820
 QY 1512 TCGGCGGCTTCGGCGCCCGCAGGGGACTCAGGGCCCTAGCCTATGCTGCGGCGCGGGA 1571
 Db 1821 TCGGCGGCTTCGGCGCCCGCAGGGGACTCAGGGCCCTAGCCTATGCTGCGGCGCGGGA 1880
 QY 1572 GCTGAGAGAGCTCTCTGTATTCTACCCAGGCGCTGGTAGCCTTCTCTGATGTGGATCT 1631
 Db 1881 GCTGAGAGAGCTCTCTGTATTCTACCCAGGCGCTGGTAGCCTTCTCTGATGTGGATCT 1940
 QY 1632 CATCTGGAAGCTTCTGAGCTGGGCGCCCTGCGGCTGAGAGCCTATGCTTCCCGCTC 1691
 Db 1941 CATCTGGAAGCTTCTGAGCTGGGCGCCCTGCGGCTGAGAGCCTATGCTTCCCGCTC 2000
 QY 1692 AGTGACCTCATCTCTGTCAGCAGCAGGGGCGCCCGAGGCTGGAGGCGACCAATTGTGT 1751
 Db 2001 AGTGACCTCATCTCTGTCAGCAGCAGGGGCGCCCGAGGCTGGAGGCGACCAATTGTGT 2060
 QY 1752 AGAGCCAGAGGGGAACCACTTTGGGAACCCCAACCTTCCATGGATGGAGAACTGCTGCT 1811
 Db 2061 AGAGCCAGAGGGGAACCACTTTGGGAACCCCAACCTTCCATGGATGGAGAACTGCTGCT 2120
 QY 1812 GAGGCGAGAGGATCTAGCCAGCAGGTGGAGGCTTGTTCAGGGGCTGGCGCTTTCAGCC 1871
 Db 2121 GAGGCGAGAGGATCTAGCCAGCAGGTGGAGGCTTGTTCAGGGGCTGGCGCTTTCAGCC 2180
 QY 1872 CTCTGGCTTGGCTTTTGTTCACACGCTG 1899

Db 2181 CTCTGGCTGGCTTTGCTTACACGCTG 2208
 RESULT 7
 AX301833
 LOCUS AX301833
 DEFINITION Sequence 12 from Patent WO0185768.
 ACCESSION AX301833
 VERSION AX301833.1 GI:17382893
 KEYWORDS
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 1
 REFERENCE
 AUTHORS Gu, W.
 TITLE Novel g-protein coupled receptors and uses therefor
 JOURNAL Patent: WO 0185768-A 12 15-NOV-2001;
 Millennium Pharmaceuticals, Inc. (US)
 FEATURES
 source Location/Qualifiers
 1..2901
 /organism="Homo sapiens"
 /mol_type="unassigned DNA"
 /db_xref="taxon:9606"
 1..>2901
 /note="unnamed protein product"
 /codon_start=1
 /protein_id="CADI3082.1"
 /db_xref="GI:17382894"
 /translation="MSPPPGLRALWLCALCASRRAGGAPQPGPPTACPAQCQD
 GIMLSADCSLGLSAPVGDLPDPLTAVLDLSMNNLTLOPLGFHHLRFLRLSGNHL
 SHIPGQAFSLYSIKILMLQNNQGLGIPABALWELPSLQSLRDANLISLVPERSFPG
 LSSRLHLDNDALEIIPVALNNLPAQALWALNLRISHIPDYAFQNLVSLVVLHLH
 NNRIOHLGTHSFEGLNLETLNLYNKLOFPVATIRLGLRELGFENNIIKAIPEKA
 FMGNPLIQTHFDNPTQFVGRSAFVLPKLHTLSLNGAMDIQEFPLDKGTTSLEILT
 LRAGIRLPLPSGMCQOLPRLVLELHNQIEBELPSLHRCOKLEIIGLOHNRIVEIGAD
 TFSQSLQALDLSMNAISIHPEAPSTLHSLVKLDLTDNQLTTLPLAGLGGMLHKL
 KGMALUSQAFSKDSFPKRLILEVYAYQCCPYGMCASFPAQSGQWENEDLHLDDESS
 KRPLGLLARQENHYDDLELQEMEDSKPHSPVQCSPTTGFPKCEYLFBSWGRIL
 AVWAILVSLVNLGLVLLTVFAGGPAPLPVFKFVNGAIGANTLITGSCGLASVDAL
 TFQFGEYGARWETGLCRATGFLVIGSEASVLLLTAAVQCSVSVSVRAVYKSPS
 LGSVRAGVGLCLAGLAAALPLASVEYGAAPLCLPYAPPEGPPALGCTVALVMNN
 SFCLVAVAGAYIKLYCDLPEDFEAVWDCAMVRHAWLIIFADGLLYCPVAFSPASML
 GLFPVTPVAVKSVLLVVLPLPACINPLLYLLFNPHRDDLRRLRPRAGDSGPLAYAAA
 GELEKSCDSTQALVAFSDVLDLLEASRPEGLTYGPPSVTLISCCQPGAPRLEG
 SHCVEPEGNHFNPNQPSMDGELLRLRAEGSTPAGGGLSGGGGFPQPSGLAFASHV"

ORIGIN
 Query Match 92.4%; Score 1754.4; DB 6; Length 2901;
 Best Local Similarity 99.9%; Pred. No. 0;
 Matches 1766; Conservative 0; Mismatches 1; Indels 1; Gaps 1;
 QY 132 AGCGGCTCCAAACACACACCGCATCTGGGAAATTTGGAGCTNGACACCTTCAGCCAGCTGAG 191
 Db 1135 ATCGGCTCCAAACACACCGCATCTGGGAAATTTGGAGCT-GACACCTTCAGCCAGCTGAG 1193
 QY 192 CTCCCTGCAAGCCCTGGAATCTTAGCTGGAACGGCATCCGCTCCATCCACCTGAGGCTT 251
 Db 1194 CTCCCTGCAAGCCCTGGAATCTTAGCTGGAACGGCATCCGCTCCATCCACCTGAGGCTT 1253
 QY 252 CTCCACCTGCACTCCCTGCTCAAGCTGAGCTGACAGCAACCGCTGACACCTGACCTGACCTGCC 311
 Db 1254 CTCCACCTGCACTCCCTGCTCAAGCTGAGCTGACAGCAACCGCTGACCTGACCTGACCTGCC 1313
 QY 312 CTCTGGCTGGAATTTGGGGCTTGTATGATCTGAAGCTCTGAAGGAACTTGTCTCTCTCCCA 371
 Db 1314 CTCTGGCTGGAATTTGGGGCTTGTATGATCTGAAGCTCTGAAGGAACTTGTCTCTCTCCCA 1373
 QY 372 GGCCTTCTCCAAAGACAGTTTCCCAAACTGAGGATCTCGAGAGTGCCTTATGCTTACCA 431
 Db 1374 GGCCTTCTCCAAAGACAGTTTCCCAAACTGAGGATCTCGAGAGTGCCTTATGCTTACCA 1433

Db	2514	TCGGCGGCTTCGGCCCGCGCAGGGGACTCAGGGCCCTAGACCTATGCTGCGCGCGGGGA	2573
Qy	1572	GCTGGGAAAGAGCTCCTGTGATTCTACCCAGGGCCCTGGTAGCCCTTCTCATGTGGATCT	1631
Db	2574	GCTGGGAAAGAGCTCCTGTGATTCTACCCAGGGCCCTGGTAGCCCTTCTCATGTGGATCT	2633
Qy	1632	CATTCTGGAAGCTTCTGAAGCTGGCGGCGCCCTTGGGCTGGAGACCTATGGCTTCCCTC	1691
Db	2634	CATTCTGGAAGCTTCTGAAGCTGGCGGCGCCCTTGGGCTGGAGACCTATGGCTTCCCTC	2693
Qy	1692	AGTGACCTCATCTCTGTGTCAGCAGCAGAGGGGCCCCCAGGCTGGAGGCGACCAATTGCT	1751
Db	2694	AGTGACCTCATCTCTGTGTCAGCAGCAGAGGGGCCCCCAGGCTGGAGGCGACCAATTGCT	2753
Qy	1752	AGAGCCAGAGGGGAACACATTTGGGAACCCCAACCCCTCCATGATGGAGAACTGCTGCT	1811
Db	2754	AGAGCCAGAGGGGAACACATTTGGGAACCCCAACCCCTCCATGATGGAGAACTGCTGCT	2813
Qy	1812	GAGGCGAGAGGATCTACGCCAGCAGGTGAGGGCTTGTCTCAGGGGGTGGCGCTTTCAGCC	1871
Db	2814	GAGGCGAGAGGATCTACGCCAGCAGGTGAGGGCTTGTCTCAGGGGGTGGCGCTTTCAGCC	2873
Qy	1872	CTCTGGGCTTGGCCTTTGCTTCACACGCG	1899
Db	2874	CTCTGGGCTTGGCCTTTGCTTCACACGCG	2901
RESULT 8			
AY358119			
LOCUS	AY358119	3325 bp mRNA linear PRI 03-OCT-2003	
DEFINITION	Homo sapiens clone DNA175959 gonadotropin receptor (UNQ6427) mRNA, complete cds.		
ACCESSION	AY358119		
VERSION	AY358119.1	GI:37181343	
KEYWORDS	FLI_CDNA:		
SOURCE	Homo sapiens (human)		
ORGANISM	Homo sapiens		
REFERENCE			
AUTHORS	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.		
	1. (bases 1 to 3325)		
	Clark,H.F., Gurney,A.L., Abaya,E., Baker,K., Baldwin,D., Brush,J., Chen,J., Chow,B., Chul,C., Crowley,C., Currell,B., Deuel,B., Dowd,P., Eaton,D., Foster,J., Grimaldi,C., Gu,Q., Haas,P.E., Heldens,S., Huang,A., Kim,H.S., Klimowski,L., Jin,Y., Johnson,S., Lee,J., Lewis,L., Liao,D., Mark,M., Robbie,E., Sanchez,C., Schoenfeld,J., Seshagiri,S., Simmons,L., Singh,J., Smith,V., Stinson,J., Vagts,A., Vanden,R., Watanabe,C., Wieand,D., Woods,K., Xie,M.H., Yansura,D., Yi,S., Yu,G., Yuan,J., Zhang,M., Zhang,Z., Goddard,A., Wood,W.I. and Godowski,P.		
TITLE	The Secreted Protein Discovery Initiative (SPDI), a Large-Scale Effort to Identify Novel Human Secreted and Transmembrane Proteins		
JOURNAL	A Bioinformatics Assessment		
PUBMED	Genome Res. 13 (10), 2265-2270 (2003)		
REFERENCE	12975309		
AUTHORS	2 (bases 1 to 3325)		
TITLE	Clark,H.F.		
JOURNAL	Direct Submission		
FEATURES	Submitted (01-AUG-2003) Department of Bioinformatics, Genentech, Inc., 1 DNA Way, South San Francisco, CA 94080, USA		
source	Location/Qualifiers		
	1. .3325		
	/organism="Homo sapiens"		
	/mol_type="mRNA"		
	/db_xref="taxon:9606"		
	/clone="DNA175959"		
	1. .3325		
	/locus tag="UNQ6427"		
	85. .2832		
	/locus tag="UNQ6427"		
	/note="PR021331"		
	/codon_start=1		
	/product="gonadotropin receptor"		
gene			
CDS			

```

/protein id="AAQ88486.1"
/db_xref="GI:37181344"
/translation="WGRPRLTLVCOVSIIISARDLSMMNTELOGLPHLRFLEELR
LSGNHLSHPQASFGSLYSKLMLQNNQLGGIPAEALWELPSLQSLRLDANILSLVP
VRSFGLSRLHMLDDNALTEIPVRLNNLPALQAMFLANRISHIPDYAFQNLTSU
VVLHNNRIQHLGTHSPGHNLETLNLNKLQEFVAIRTLRQLQELGFHNNNIK
AIPKAFNGPILLOTIHPYDNP IQFVRSAPFQPLKHLTSLINGAMDIQEPDLKGT
SLEILTFRAGRLTLPQSMCOOLPRLRVLELSHNOIBELPSLHRCOKLEIGLOHNR
WEIGADTFPSQLSALDLSNNAIRSHPPEARSTLHSLVKLDLTDNOLTTPLAGLG
LMLHKLGNLALSAFQSKDPSFKRLILEVYAYOCCPYGMCASFQKSGQWEABDLHL
DDESSRGLLGLLRASNDYDQDLDELQLEMEDSKPHPSVQCSPTFPFKPBYLLE
SWGRLAVVAILSVLSCNGLVLLTVFAGGAPLPVPKVVVGAIGANTLGISGCL
ASVDALTQWQSEYGARWETGLGRATFLAVLSEASVILLTLAAVQCSVSVSCVRA
YKSPSLGSVRAVLGCLALAAALPLASVGEVGSPLCLPVAPGEPQALAGFTV
ALVMNNSFCFLVAGAYIKLYCDLPRGDFEAVMDCMVRHVAMLIIFADGLLYCPVAF
SPASMLGUPFPVTPERAKSVLVLPLPACLNPLLLPNPHFDRLDLRPRGSGP
LAYAAAGLEKSDSTQALVAFSDVDLILEASRPPGLETYGFPSVTLISCOQPG
APRLGSHCVPEGNHFGNPQPSMDGELLRLRAEGSTPAGGSLSGGGGFQPSGLAFAS
H"

```

ORIGIN

Query Match 92.4%; Score 1754.4; DB 9; Length 3325;
 Best Local Similarity 99.9%; Pred. No. 0;
 Matches 1766; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

Qy	132	AGCGSCCTCCAAACACACCGCATCTGGGAAATTGGAGCTNGACACCTTTCAGCCAGCTGAG	191
Db	1063	ATCGGCCCTCCAAACACACCGCATCTGGGAAATTGGAGCT - GACACCTTCAGCCAGCTGAG	1121
Qy	192	CTCCCTGCAAGCCCTGGATCTTAGCTGGAACGCCATCCGGTCCATCCACCCCTGAGGCCCTT	251
Db	1122	CTCCCTGCAAGCCCTGGATCTTAGCTGGAACGCCATCCGGTCCATCCACCCCTGAGGCCCTT	1181
Qy	252	CTCCACCTCGACCTCCCTGGTCAAGCTGGACCTGACAGAACACCGAGCTGACCACTGGCC	311
Db	1182	CTCCACCTCGACCTCCCTGGTCAAGCTGGACCTGACAGAACACCGAGCTGACCACTGGCC	1241
Qy	312	CCTGGCTGACCTTGGGGCTTGATGCATCTGAAGCTCAAGGGAACCTTGCTCTCTCCCA	371
Db	1242	CCTGGCTGACCTTGGGGCTTGATGCATCTGAAGCTCAAGGGAACCTTGCTCTCTCCCA	1301
Qy	372	GGCCTTCTCCAAAGACAGTTTCCCAAAACTCAGAGATCTCGAGGTGCTTATGCTACCA	431
Db	1302	GGCCTTCTCCAAAGACAGTTTCCCAAAACTCAGAGATCTCGAGGTGCTTATGCTACCA	1361
Qy	432	GTGCTGTCCCTATGGGATGTGTGCAGCTTCTTCAAGGCCCTCTGGGAGTGGGAGGTGA	491
Db	1362	GTGCTGTCCCTATGGGATGTGTGCAGCTTCTTCAAGGCCCTCTGGGAGTGGGAGGTGA	1421
Qy	492	AGACCTTCACCTTGATGATGAGGAGTCTTCAAAAGGCCCTCGGCCCTCTTCCGACACA	551
Db	1422	AGACCTTCACCTTGATGATGAGGAGTCTTCAAAAGGCCCTCGGCCCTCTTCCGACACA	1481
Qy	552	AGCAGAGAACCACTATGACAGGACCTGGATGAGCTCCAGCTGGAGATGGAGGACTCAAA	611
Db	1482	AGCAGAGAACCACTATGACAGGACCTGGATGAGCTCCAGCTGGAGATGGAGGACTCAAA	1541
Qy	612	GCCACACCCAGTGTCCAGTGTAGCCCTTATCCAGGCCCTTCAAGGCCCTGTGAGTACCT	671
Db	1542	GCCACACCCAGTGTCCAGTGTAGCCCTTATCCAGGCCCTTCAAGGCCCTGTGAGTACCT	1601
Qy	672	CTTTGAAAGCTGGGGCATCCGCTGGCGTGTGGGCCATCGTGTGCTCTCCGTGCTCTG	731
Db	1602	CTTTGAAAGCTGGGGCATCCGCTGGCGTGTGGGCCATCGTGTGCTCTCCGTGCTCTG	1661
Qy	732	CAATGGACTGTGTGCTGACCGTGTTCGCTGGCGGGCTCCGCCCTGCGCCCGGTCAA	791
Db	1662	CAATGGACTGTGTGCTGACCGTGTTCGCTGGCGGGCTCCGCCCTGCGCCCGGTCAA	1721
Qy	792	GTTTGTGGTAGGTGCGATTGACGGCGCCAAACCTTGACTGGCATTTCTGTGGCCTTCT	851
Db	1722	GTTTGTGGTAGGTGCGATTGACGGCGCCAAACCTTGACTGGCATTTCTGTGGCCTTCT	1781

Qy	852	AGCCTCAGTCGATGCCCTGACCTTTGGTTCAGTTTCTCTGAGTACGAGAGCCCTCGGAGAC	911
Db	1782	AGCCTCAGTCGATGCCCTGACCTTTGGTTCAGTTTCTCTGAGTACGAGAGCCCTCGGAGAC	1841
Qy	912	GGGCTAGGCTGCGCGGCGCACTGCTCTTCTGGCAGTACTTGGGTTCGAGGCAATCGGTGCT	971
Db	1842	GGGCTAGGCTGCGCGGCGCACTGCTCTTCTGGCAGTACTTGGGTTCGAGGCAATCGGTGCT	1901
Qy	972	GCTGCTCACTCTGGCGCAGTGCAGTGCAGCGTCTCCGTCTCCCTGTGTCGCGGCTATGG	1031
Db	1902	GCTGCTCACTCTGGCGCAGTGCAGTGCAGCGTCTCCGTCTCCCTGTGTCGCGGCTATGG	1961
Qy	1032	GAACTCCCTCCCTGGGCGAGCGTTTCGAGCAGGGGTCTTAGGCTGCTGGCACTGGCAGG	1091
Db	1962	GAACTCCCTCCCTGGGCGAGCGTTTCGAGCAGGGGTCTTAGGCTGCTGGCACTGGCAGG	2021
Qy	1092	GCTGGCGCGCGCACTGCCCCCTGGGCTCAGTGGGAGAAATACGGGGCTTCCCCACTCTGCT	1151
Db	2022	GCTGGCGCGCGCACTGCCCCCTGGGCTCAGTGGGAGAAATACGGGGCTTCCCCACTCTGCT	2081
Qy	1152	GCCCTACGCGCCACCTGAGGGTCAGCCAGCAGCCCTGGGCTTACCGTGGCCCTGGTGAT	1211
Db	2082	GCCCTACGCGCCACCTGAGGGTCAGCCAGCAGCCCTGGGCTTACCGTGGCCCTGGTGAT	2141
Qy	1212	GATGAACCTCTCTGTTTCTGTCGTGGCGGTGCTCATCAAACTGTACTGTGACCT	1271
Db	2142	GATGAACCTCTCTGTTTCTGTCGTGGCGGTGCTCATCAAACTGTACTGTGACCT	2201
Qy	1272	GCGCGGGGCGACTTTGAGGCCGTGTGGGACTGCGCCATCGTGGAGCACTGGGCTGGCT	1331
Db	2202	GCGCGGGGCGACTTTGAGGCCGTGTGGGACTGCGCCATCGTGGAGCACTGGGCTGGCT	2261
Qy	1332	CATCTTCGACAGCGGGCTCTTACTGTCCGTGGGCTTCTCTCAGCTTGGCTTCATGCT	1391
Db	2262	CATCTTCGACAGCGGGCTCTTACTGTCCGTGGGCTTCTCTCAGCTTGGCTTCATGCT	2321
Qy	1392	GGGCTCTTCCCTGTACGCGCCGAGGCGCTCAAGTCTGCTCTGCTGTGTGCTGCGCCCT	1451
Db	2322	GGGCTCTTCCCTGTACGCGCCGAGGCGCTCAAGTCTGCTCTGCTGTGTGCTGCGCCCT	2381
Qy	1452	GCTGCTGTGCTCAACCCACTGTGCTACCTGCTCTTCAACCCCCACCTTCCGGGATGACCT	1511
Db	2382	GCTGCTGTGCTCAACCCACTGTGCTACCTGCTCTTCAACCCCCACCTTCCGGGATGACCT	2441
Qy	1512	TGCGGCGCTTTCGCGCCCGCGCAGGAGACTCAGGGGCCCTTAGCTATGCTGCGGCGCGGGA	1571
Db	2442	TCGGCGGCTTTCGCGCCCGCGCAGGAGACTCAGGGGCCCTTAGCTATGCTGCGGCGCGGGA	2501
Qy	1572	GCTGGAGAGAGCTCTCTGTGATTTCTACCCAGGCGCTGGTAGCTTCTCTGATGTGGATCT	1631
Db	2502	GCTGGAGAGAGCTCTCTGTGATTTCTACCCAGGCGCTGGTAGCTTCTCTGATGTGGATCT	2561
Qy	1632	CATTCTGGAAGCTTCTGAAGCTGGGCGGCCCTTGGGCTGGAGACCTATGGCTTCCCTC	1691
Db	2562	CATTCTGGAAGCTTCTGAAGCTGGGCGGCCCTTGGGCTGGAGACCTATGGCTTCCCTC	2621
Qy	1692	AGTGAACCTCATCTCTGTGACAGCCAGGGGCGCCCGAGGCTGGAGGCGACCAATTGTGT	1751
Db	2622	AGTGAACCTCATCTCTGTGACAGCCAGGGGCGCCCGAGGCTGGAGGCGACCAATTGTGT	2681
Qy	1752	AGAGCCAGAGGGGAAACCACTTTTGGGAAACCCCAACCCCTCCATGATGGAGAACTGCTGCT	1811
Db	2682	AGAGCCAGAGGGGAAACCACTTTTGGGAAACCCCAACCCCTCCATGATGGAGAACTGCTGCT	2741
Qy	1812	GAGGCGAGAGGGGACTTACGCCAGCAGGTGAGGCTTGTCTAGGGGGTGGCGGCTTTCAGCC	1871
Db	2742	GAGGCGAGAGGGGACTTACGCCAGCAGGTGAGGCTTGTCTAGGGGGTGGCGGCTTTCAGCC	2801
Qy	1872	CTCTGGCTTGGCCTTTGCTTCAACAGTG	1899
Db	2802	CTCTGGCTTGGCCTTTGCTTCAACAGTG	2829

```
RESULT 9
AX301831
LOCUS AX301831 3492 bp DNA linear PAT 30-NOV-2001
DEFINITION Sequence 10 from Patent WO0185768.
ACCESSION AX301831
VERSION AX301831.1 GI:17382891
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1
REFERENCE
AUTHORS Gu.W.
TITLE Novel g-protein coupled receptors and uses therefor
JOURNAL Patent: WO 0185768-A 10 15-NOV-2001;
Millennium Pharmaceuticals, Inc. (US)
FEATURES
Source
1..3492 Location/Qualifiers
    /organism="Homo sapiens"
    /mol_type="unassigned DNA"
    /db_xref="taxon:9606"
    104..3007
    /note="unnamed protein product"
    /codon_start=1
    /protein_id="CAD13081.1"
    /db_xref="GI:17382892"
    /translation="MPSPPGRLMCLCAALCASRRAGAPQPGPTACAPACPCQED
    SHILSADCSGLGSAVPGDLPDLPATAYLDLSMNNLTPELQGLFHLRLFLLELRLSGNHL
    SHIPQAPSGIYSKILMLQNNOLGGIPAEALWELPSLQSLRLDANLILSVPSRSEF
    LSSRLHLWDNALTEIPVRALNLPALQANTLALNRISHIPDYAFQNLTSLVVHLH
    NNRLOHGLTFEGUHNLETDLYNKLQEPFAIRITUGRLQELGFHNNKAIPEKA
    FMGNPLQTHIFHYNIQFVGRSAFYQLPLKHLTSLNGAMDIQEPFDLJKGTTLSLEIL
    LTRAGIRLLPSGMCQQLPRLRVLELSHQIIEBLSLHRCQKLEIGLQHNRIWIGAD
    TFSOLSQALDLSNARLSRHPAFSLHSLVKLDLTDNOLTTPLAGLGLMLHLK
    KGNLALSAFSGKSPKRLILEVPAYOCYVGMCAFPFKASGOWEABDLDRDEESS
    KRPLGLLARQENHYDQDLDELQLEMEDSKHPSPVQCSPTPGPKPCYLPESGIRL
    AVMAILVSLVNGLVLLTVFAGGPAPLPVVKVVGATAGANTLTGISCGLLASVDAL
    TFGQFSEYGAEMETGLGRATGFLAVLGSEASVILLTLAAVQCSVSVSCVAYKGSPS
    LGSVRAGVLGCLAGLAAALPLASVGEYGASPLCLPYAPPEGQPAALGFTVALVMNN
    SFCFLVAGAVIKLYCDLPRGDFSAVMDCAVNRHVAVLIFADGLLYCFVAFLSFASML
    GLFPVTPAVKSVLLVLPACLNPLLYLLAFNHFDRDLRLPRAGDSGLPAAVA
    GELEKSSCDSTQALVARSVDVLIILEASGRPPGLETYGPPSVTLISQQQGPAPLEG
    SHCPEPENGHNFPQPSMDGELLURABGSTFAGGSLGGGFGQPSGLAFASHV"
ORIGIN
Query Match 92.4%; Score 1754.4; DB 6; Length 3492;
Best Local Similarity 99.9%; Pred No. 0;
Matches 1766; Conservative 0; Mismatches 1; Indels 1; Gaps 1;
Qy 132 AGCGGCTCCAAACACACCGCATCTGGGAAATTGGAGCTNGACACCTTCAGCCAGCTGAG 191
Db 1238 ATCGGCTCCAAACACACCGCATCTGGGAAATTGGAGCT-GACACCTTCAGCCAGCTGAG 1296
Qy 192 CTCCTGCAAGCCCTGGATCTTAGCTGAAACGCCATCCGGTCCATCCACCTGAGGCTT 251
Db 1297 CTCCTGCAAGCCCTGGATCTTAGCTGAAACGCCATCCGGTCCATCCACCTGAGGCTT 1356
Qy 252 CTCACCTCGACCTCCCTGGTCAAGCTGACCTGACACACACACACACCTGCTCTCTCCA 311
Db 1357 CTCACCTCGACCTCCCTGGTCAAGCTGACCTGACACACACACACACCTGCTCTCTCCA 1416
Qy 312 CTCGCTGGACTTGGGGGCTTGATGCATCTGAAGCTCAAAGGAAACCTTGCTCTCTCCA 371
Db 1417 CTCGCTGGACTTGGGGGCTTGATGCATCTGAAGCTCAAAGGAAACCTTGCTCTCTCCA 1476
Qy 372 GGCCTTCTCCAAAGACAGTTTCCCAAACTGAGGATCTCGAGGTGCCTTATGCCTACA 431
Db 1477 GGCCTTCTCCAAAGACAGTTTCCCAAACTGAGGATCTCGAGGTGCCTTATGCCTACA 1536
Qy 432 GTGCTGCTCCATGGAGTGTGGCAGCTTCTTCAAGGCTCTGGGAGTGGGAGGCTGA 491
Db 1537 GTGCTGCTCCATGGAGTGTGGCAGCTTCTTCAAGGCTCTGGGAGTGGGAGGCTGA 1596
```


Db 2677 GCTGGAGAGAGCTCCTGTGATTTACCCAGGCCCTGTGTAGCCTTCTCTGATGTGGATCT 2736
Qy 1632 CATTCTGGAAGCTTCTGAAGCTGGGCGGCCCTCTGGGCTGGAGACCTATGCGTTCCTCCCTC 1691
Db 2737 CATTCTGGAAGCTTCTGAAGCTGGGCGGCCCTCTGGGCTGGAGACCTATGCGTTCCTCCCTC 2796
Qy 1692 AGTGACCTCATCTCTCTGTCAGCAGCCAGGGGCCCCCAGGCTGGAGGGCAGCCATTGTGT 1751
Db 2797 AGTGACCTCATCTCTCTGTCAGCAGCCAGGGGCCCCCAGGCTGGAGGGCAGCCATTGTGT 2856
Qy 1752 AGAGCCAGAGGGGAACCACTTTGGGAACCCCAACCCCTCCATGATGAGAACTGCTGCT 1811
Db 2857 AGAGCCAGAGGGGAACCACTTTGGGAACCCCAACCCCTCCATGATGAGAACTGCTGCT 2916
Qy 1812 GAGGCGAGAGGATCTAGCCAGCAGGTGGAGGCTTGTTCAGGGGTGGCGCTTTCAGCC 1871
Db 2917 GAGGCGAGAGGATCTAGCCAGCAGGTGGAGGCTTGTTCAGGGGTGGCGCTTTCAGCC 2976
Qy 1872 CTCTGGCTTGGCTTGTCTTACACGTG 1899
Db 2977 CTCTGGCTTGGCTTGTCTTACACGTG 3004

RESULT 10
LOCUS AX549295 3119 bp DNA linear PAT 26-NOV-2002
DEFINITION Sequence 580 from Patent WO02061087.
ACCESSION AX549295
VERSION AX549295.1 GI:25813958
KEYWORDS .
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1
REFERENCE 1
AUTHORS Burmer,G.C., Roush,C.L. and Brown,J.P.
TITLE Antigenic peptides, such as for G protein-coupled receptors (GPCRs), antibodies thereto, and systems for identifying such antigenic peptides
JOURNAL Patent: WO 02061087-A 580 08-AUG-2002;
Lifespan Biosciences, Inc. (US)
FEATURES
source 1. 3119
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

Query Match 92.3%; Score 1752.8; DB 6; Length 3119;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 1765; Conservative 0; Mismatches 2; Indels 1; Gaps 1;

Qy 132 AGCGGCTCCAAACAAACCGCATCTGGGAATTTGGAGCTNGACACCTTTCAGCCAGCTGAG 191
Db 1020 ATCGGCTCCAAACAAACCGCATCTGGGAATTTGGAGCT-GACACCTTCAGCCAGCTGAG 1078
Qy 192 CTCCTGCAAGCCCTGATCTTAGTGGAACGCCATCCGGTCCATCCACCTGAGGCTT 251
Db 1079 CTCCTGCAAGCCCTGATCTTAGTGGAACGCCATCCGGTCCATCCACCTGAGGCTT 1138
Qy 252 CTCACACCTGACATCCCTGGTCAAGCTGGACCTGACAGAACACAGCTGACCACTGCC 311
Db 1139 CTCACACCTGACATCCCTGGTCAAGCTGGACCTGACAGAACACAGCTGACCACTGCC 1198
Qy 312 CCTGGCTGGACTTGGGGCTTGTATGCATCTGAAGCTCAAAAGGGAACCTTGTCTCTCCCA 371
Db 1199 CCTGGCTGGACTTGGGGCTTGTATGCATCTGAAGCTCAAAAGGGAACCTTGTCTCTCCCA 1258
Qy 372 GGCTTCTCCAAAGACAGTTTCCAAAACTGAGGATCTTGAGGTGCTTATGCTTACCA 431
Db 1259 GGCTTCTCCAAAGACAGTTTCCCAAACTGAGGATCTTGAGGTGCTTATGCTTACCA 1318

Qy 432 GTGCTGTCCCTATATGGGATGTGTGCCAGCTTCTTCAAGGCCCTCTGGGAGTGGGAGGCTGA 491
Db 1319 GTGCTGTCCCTATATGGGATGTGTGCCAGCTTCTTCAAGGCCCTCTGGGAGTGGGAGGCTGA 1378
Qy 492 AGACCTTCACCTTGATGATGAGGAGTCTTCAAAAGGCCCTCTGGGCTCTCTTGCAGACA 551
Db 1379 AGACCTTCACCTTGATGATGAGGAGTCTTCAAAAGGCCCTCTGGGCTCTCTTGCAGACA 1438
Qy 552 AGCAGAGAACACATATCACAGGACCTTGGATGAGCTCCAGCTGAGATGGAGGACTCAAA 611
Db 1439 AGCAGAGAACACATATCACAGGACCTTGGATGAGCTCCAGCTGAGATGGAGGACTCAAA 1498
Qy 612 GCCACACCCAGTGTCCAGTGTAGCCCTACTTCCAGGCCCTTCAAGCCCTGTGAGTACCT 671
Db 1499 GCCACACCCAGTGTCCAGTGTAGCCCTACTTCCAGGCCCTTCAAGCCCTGTGAGTACCT 1558
Qy 672 CTTTGAAGCTGGGGCATCCGCTTGGCGCTGTGGGCCCATCGTGTGTGCTCTCCGTGCTGTG 731
Db 1559 CTTTGAAGCTGGGGCATCCGCTTGGCGCTGTGGGCCCATCGTGTGTGCTCTCCGTGCTGTG 1618
Qy 732 CAATGGACTGGTGTGCTGACCGTGTTCGCTGGCGGGCTGCCCTGCCCTGCCCTGGCTCAA 791
Db 1619 CAATGGACTGGTGTGCTGACCGTGTTCGCTGGCGGGCTGCCCTGCCCTGGCTCAA 1678
Qy 792 GTTTGTGTAGTGCAGATTGCAGCGCCAAACACCTTGACTGGCATTTCTGTGGCCTTCT 851
Db 1679 GTTTGTGTAGTGCAGATTGCAGCGCGCAACACCTTGACTGGCATTTCTGTGGCCTTCT 1738
Qy 852 AGCCTCAGTCGATGCCCTGACCTTTTGGTTCAGTTCTCTGAGTACGGAGCCGCTGGAGAC 911
Db 1739 AGCCTCAGTCGATGCCCTGACCTTTTGGTTCAGTTCTCTGAGTACGGAGCCGCTGGAGAC 1798
Qy 912 GGGGCTAGGCTGCGGGCCACTGGCTTCTCGGAGTACTTTGGGTCGAGGACATCGGTGCT 971
Db 1799 GGGGCTAGGCTGCGGGCCACTGGCTTCTCGGAGTACTTTGGGTCGAGGACATCGGTGCT 1858
Qy 972 GCTGCTCACTCTGGCCAGTCAGTACAGCTTCTCGTCTCTGCTGTCTCGGGGCTATGG 1031
Db 1859 GCTGCTCACTCTGGCCGACAGTCAGTACAGGCTTCTCGTCTCTGCTGTCTCGGGGCTATGG 1918
Qy 1032 GAAGTCCCTTCTCTGGGAGCTTTCAGCAGGGGCTCTAGGCTGCTGCGACATGGCAGG 1091
Db 1919 GAAGTCCCTTCTCTGGGAGCTTTCAGCAGGGGCTCTAGGCTGCTGCGACATGGCAGG 1978
Qy 1092 GCTGGCGCGGACATGCGGCTGCGCTCAGTGGGAGATAGGGGCTTCCCACTCTGCCT 1151
Db 1979 GCTGGCGCGGACATGCGGCTTGGCTTCTCAGTGGGAGATAGGGGCTTCCCACTCTGCCT 2038
Qy 1152 GCCCTACGCGCCACCTGAGGCTCAGCAGCAGCCCTGGGCTTCAACCGTGGCCCTGTGTGAT 1211
Db 2039 GCCCTACGCGCCACCTGAGGCTCAGCAGCAGCCCTGGGCTTCAACCGTGGCCCTGTGTGAT 2098
Qy 1212 GATGAATCTCTTCTGTTTCTGTGTCGTGGCGGTGCTACATCAAACTGTGTGACCT 1271
Db 2099 GATGAATCTCTTCTGTTTCTGTGTCGTGGCGGTGCTACATCAAACTGTGTGACCT 2158
Qy 1272 GCGCGGGGCGACTTTGAGGCCGTGTGGGACTGCGGCATGCTGAGGACATGCGGCTGGCT 1331
Db 2159 GCGCGGGGCGACTTTGAGGCCGTGTGGGACTGCGGCATGCTGAGGACATGCGGCTGGCT 2218
Qy 1332 CATCTTCGACAGAGGGCTCTCTACTGTCCCGTGGCTTCTCAGCTTCCCTCCATGCT 1391
Db 2219 CATCTTCGACAGAGGGCTCTCTACTGTCCCGTGGCTTCTCAGCTTCCCTCCATGCT 2278
Qy 1392 GGGCTCTTCTCTGTCAAGCCGAGGCGCTCAAGTGTGTCTCTGCTGGTGGTGGCTCCCT 1451
Db 2279 GGGCTCTTCTCTGTCAAGCCGAGGCGCTCAAGTGTGTCTCTGCTGGTGGTGGCTCCCT 2338
Qy 1452 GCCTGCTGCTCAACCCACTGCTACTGCTTTCACACCCCACTTCCGGGATGACCT 1511
Db 2339 GCCTGCTGCTCAACCCACTGCTACTGCTTTCACACCCCACTTCCGGGATGACCT 2398
Qy 1512 TCGGCGGCTTTCGGCCCGCGGACATGAGGGCCCTAGCCTATAGCTGCGGCGGGGA 1571

Db 2399 TCGGCGGCTTCGGCCCGCGCAGGCGACTCAGGGCCCTTAGCCTATGCTGCGCGCGGGA 2458
Qy 1572 GCTGAGAAAGAGCTCTGATGATTTACCCAGGCGCTTGTGATGCTTCTCTGATGCGATCT 1631
Db 2459 GCTGAGAAAGAGCTCTGATGATTTACCCAGGCGCTTGTGATGCTTCTCTGATGCGATCT 2518
Qy 1632 CATCTGGAAGCTTCTGAAGCTGGCGCGCCCTCGGCTGGAGACTATGGCTTCCCTC 1691
Db 2519 CATCTGGAAGCTTCTGAAGCTGGCGCGCCCTCGGCTGGAGACTATGGCTTCCCTC 2578
Qy 1692 AGTGACCTCATCTCTGTCAGCAGCAGGCGCGCCCGAGGCTGGAGGCGAGCATGTGT 1751
Db 2579 AGTGACCTCATCTCTGTCAGCAGCAGGCGCGCCCGAGGCTGGAGGCGAGCATGTGT 2638
Qy 1752 AGAGCAGAGGGAACCACTTGGGAACCCCAACCTCCATGGATGAGAACTGCTGCT 1811
Db 2639 AGAGCAGAGGGAACCACTTGGGAACCCCAACCTCCATGGATGAGAACTGCTGCT 2698
Qy 1812 GAGGCGAGAGGATCTAGCCAGCAGGTGGAGGCTTGTGAGGGGTGGCGGCTTTCAGCC 1871
Db 2699 GAGGCGAGAGGATCTAGCCAGCAGGTGGAGGCTTGTGAGGGGTGGCGGCTTTCAGCC 2758
Qy 1872 CTCTGGCTTGGCTTGTTCACAGTG 1899
Db 2759 CTCTGGCTTGGCTTGTTCACAGTG 2786

RESULT 11
AB049405
LOCUS
DEFINITION Homo sapiens VTS20631 mRNA, 3119 bp mRNA linear PRI 11-OCT-2002
partial cds.
ACCESSION AB049405
VERSION AB049405.1 GI:13447609
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1
REFERENCE
AUTHORS Okaze,H., Hayashi,A., Kozuma,S. and Saito,T.
TITLE a member of g-protein coupled receptor family
JOURNAL Published Only in Database (2001)
REFERENCE 2 (bases 1 to 3119)
AUTHORS Okaze,H., Hayashi,A., Kozuma,S. and Saito,T.
TITLE Direct Submission
JOURNAL Submitted (27-SEP-2000) Toshiyuki Saito, National Institute of
Radiological Sciences, Genome Research Group; Anagawa 4-9-1,
Chiba, Inage-ku, Chiba 263-8555, Japan (E-mail:t_saito@nirs.go.jp,
Tel:81-43-206-3135, Fax:81-43-251-9818)
FEATURES
source
1..3119
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/chromosome="1"
/cell_type="pooled"
1..3119
/gene="VTS20631"
<1..2789
/gene="VTS20631"
/note="g-protein coupled receptor family"
/codon_start=3
/product="VTS20631"
/protein_id="BAB39854.1"
/db_xref="GI:13447610"
/translation="HCQDGMLSADCSBELGLSVPGDLDPLTAYLDLSMNNLTLOP
GLFHLRLLELRSLGSHLSHPGOAFSGLSLKLMLQNNLGLGIPAEALWELPSLO
SIRLDANLISLVSPERSFGLSSRLHLDLNDALTEIPYRALNLPALQAMTLANRIS
HIPDYAFQNLTSVVLHNNRIQHLGTHSEFGLHNLETLDLNNKLPQEFVAIRLIG
RLQELGFNNIKAIPEKAFMGNPLLOTIHFYDNPQFVRSAPQYLPKLHLSLNG
MDIQEFPDLKGTTSLEILTLAGIRLLPSGWCQQLPRLRLVLELSHNIQIEPLSIHRC

ORIGIN

Query Match 92.3%; Score 1752.8; DB 9; Length 3119;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 1765; Conservative 0; Mismatches 2; Indels 1; Gaps 1;
Qy 132 AGCGCCCTCCAAACACACCGCATCTGGGAAATTTGGAGCTNGACACCTTCCAGCAGCTGAG 191
Db 1020 ATCGGCTTCCAAACACACCGCATCTGGGAAATTTGGAGCT-GACACCTTTCAGCAGCTGAG 1078
Qy 192 CTCCCTGCAAGCCCTGGATCTTAGCTGGAACGCCATCCGGTCCATCCACCTGAGGCGCTT 251
Db 1079 CTCCCTGCAAGCCCTGGATCTTAGCTGGAACGCCATCCGGTCCATCCACCTGAGGCGCTT 1138
Qy 252 CTCACACCTGCACTCCCTGCTCAAGCTGGAACCTGACAGACAAACAGCTGACACACACTGCC 311
Db 1139 CTCACACCTGCACTCCCTGCTCAAGCTGGAACCTGACAGACAAACAGCTGACACACTGCC 1198
Qy 312 CTGCTGCTGCACTTGGGGCTTGGATGATCTGAAGCTCAAGAGGAACTTGTCTCTCTCCCA 371
Db 1199 CTTGGCTGCACTTGGGGCTTGGATGATCTGAAGCTCAAGAGGAACTTGTCTCTCTCCCA 1258
Qy 372 GGCCTTCTCCAAAGACAGTTTCCAAACACTCAGAGATCTCGAGGTGCTTATGCTTACCA 431
Db 1259 GGCCTTCTCCAAAGACAGTTTCCAAACACTCAGAGATCTCGAGGTGCTTATGCTTACCA 1318
Qy 432 GTGCTGTCTTATGGGATGTGTGCGAGCTTCTTCAAGGCTCTTGGGAGTGGGAGGCTGA 491
Db 1319 GTGCTGTCTTATGGGATGTGTGCGAGCTTCTTCAAGGCTCTTGGGAGTGGGAGGCTGA 1378
Qy 492 AGACCTTCACTTGTATGATGAGGATCTTCAAAAGGCGCTTGGGCTCTCTTGCAGACA 551
Db 1379 AGACCTTCACTTGTATGATGAGGATCTTCAAAAGGCGCTTGGGCTCTCTTGCAGACA 1438
Qy 552 AGCAGAGAACCACTATGACACAGGACCTGGATGAGCTCCAGCTGGAGATGGAGGACTCAA 611
Db 1439 AGCAGAGAACCACTATGACACAGGACCTGGATGAGCTCCAGCTGGAGATGGAGGACTCAA 1498
Qy 612 GCCACACCCCAAGTGTCCAGCTAGCCCTTACTCCAGGCGCTTCAAGCCCTGTGAGTACT 671
Db 1499 GCCACACCCCAAGTGTCCAGCTAGCCCTTACTCCAGGCGCTTCAAGCCCTGTGAGTACT 1558
Qy 672 CTTTGAAGCTGGGCGCATCCGCTGGCGGTGTGGGCGCATCGTGTGCTCTCCGTGCTCTG 731
Db 1559 CTTTGAAGCTGGGCGCATCCGCTGGCGGTGTGGGCGCATCGTGTGCTCTCCGTGCTCTG 1618
Qy 732 CAATGGACTGGTGTGCTGACCGGTGTGGTGGGCGGCTGCGGCGGCTGCGGCGGCTCAA 791
Db 1619 CAATGGACTGGTGTGCTGACCGGTGTGGTGGGCGGCTGCGGCGGCTGCGGCGGCTCAA 1678
Qy 792 GTTTGTGTAGTGTGCGATTTGAGGCGGCAACACTTGTGAGTGGCATTTCTGTGGCTTCT 851
Db 1679 GTTTGTGTAGTGTGCGATTTGAGGCGGCAACACTTGTGAGTGGCATTTCTGTGGCTTCT 1738
Qy 852 AGCTCAGTGCATGCCCTGACCTTTTGTGTCAGTTCTCTGAGTACGAGAGGCGGCTGGAGAC 911
Db 1739 AGCTCAGTGCATGCCCTGACCTTTTGTGTCAGTTCTCTGAGTACGAGAGGCGGCTGGAGAC 1798
Qy 912 GGGCTAGGCTGCGGCGGCGGCTTCTGCGAGTACTTGGGTGGAGGCGGCTGGGTGCT 971
Db 1799 GGGCTAGGCTGCGGCGGCGGCTTCTGCGAGTACTTGGGTGGAGGCGGCTGGGTGCT 1858

Db 1377 GTTGTGTAGTGGATTTGACGGCGCAACACACCTTGACTGGCAATTTCTGTGGCTTCT 1436
QY 852 AGCCTCAGTCAGTGCCTGACCTTTGGTCAAGTTCTCTGAGTACGAGCCCGCTGGAGAC 911
Db 1437 AGCCTCAGTCAATGCGCTGACCTTTGGTCAAGTTCTCTGAGTACGAGCCCGCTGGAGAC 1496
QY 912 GGGGCTAGGCTGCGGGGCACTGGCTTCTTGGCAGTACTTGGTGGAGGCAATCGGTGCT 971
Db 1497 GGGGCTAGGCTGCGGGGCACTGGCTTCTTGGCAGTACTTGGTGGAGGCAATCGGTGCT 1556
QY 972 GCTGCTCACTCTGGCGCAGTGCAGTGCAGCGCTCTCCGTCTCTCTGTGTCGGGCGCTATGG 1031
Db 1557 GCTGCTCACTCTGGCGCAGTGCAGTGCAGCGCTCTCCGTCTCTCTGTGTCGGGCGCTATGG 1616
QY 1032 GAAGTCCCTCTCTGGGAGAGGTTTCGAGCAGGGGTCTTAGCTGCCCTGGCACTGGCAGG 1091
Db 1617 GAAGTCCCTCTCTGGGAGAGGTTTCGAGCAGGGGTCTTAGCTGCCCTGGCACTGGCAGG 1676
QY 1092 GCTGGCGCGCACTGCCCTGGCTCAGTGGGAGAAACGGGGGCTCCCACTCTGCCT 1151
Db 1677 GCTGGCGCGCACTGCCCTGGCTCAGTGGGAGAAACGGGGGCTCCCACTCTGCCT 1736
QY 1152 GCGCTACGCGCACCTGAGGGTTCAGCAGCAGCCTGGGCTTCACCGTGGGCGCTGGTGAT 1211
Db 1737 GCGCTACGCGCACCTGAGGGTTCAGCAGCAGCCTGGGCTTCACCGTGGGCGCTGGTGAT 1796
QY 1212 GATGAATCTCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1271
Db 1797 GATGAATCTCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1856
QY 1272 GCGCGGGGCGACTTTGAGGCGCTGCTGAGGCTGCGCCATGCTGAGGCACTGGGCTGGCT 1331
Db 1857 GCGCGGGGCGACTTTGAGGCGCTGCTGAGGCTGCGCCATGCTGAGGCACTGGGCTGGCT 1916
QY 1332 CATCTTCGACAGCGGCTCTCTACTGCTCCGCTGGCTTCTCTGCTGCTGCTGCTGCTGCT 1391
Db 1917 CATCTTCGACAGCGGCTCTCTACTGCTCCGCTGGCTTCTCTGCTGCTGCTGCTGCT 1976
QY 1392 GGGCTCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1451
Db 1977 GGGCTCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2036
QY 1452 GCGTGCCTGCTCAACCCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1511
Db 2037 GCGTGCCTGCTCAACCCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2096
QY 1512 TCGGCGGCTTCGGGCGCGCAGGGGACTCAGGGCGGCTGAGGCTGCTGCTGCTGCTGCTGCT 1571
Db 2097 TCGGCGGCTTCGGGCGCGCAGGGGACTCAGGGCGGCTGAGGCTGCTGCTGCTGCTGCTGCT 2156
QY 1572 GCTGAGAGAGCTCTCTGTGATTTACCCAGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1631
Db 2157 GCTGAGAGAGCTCTCTGTGATTTACCCAGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2216
QY 1632 CATCTGGAAGCTTCTGAGCTGGGCGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1691
Db 2217 CATCTGGAAGCTTCTGAGCTGGGCGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2276
QY 1692 AGTGACCTCATCTCTGCTGAGCAGCAGGGGCGGCGGCTGAGGCTGAGGCGACCCATTTGCT 1751
Db 2277 AGTGACCTCATCTCTGCTGAGCAGCAGGGGCGGCGGCTGAGGCTGAGGCGACCCATTTGCT 2336
QY 1752 AGAGCAGAGGGGAAACCACTTTGGGAAACCCCAACCTTCATGATGAGAACTGCTGCT 1811
Db 2337 AGAGCAGAGGGGAAACCACTTTGGGAAACCCCAACCTTCATGATGAGAACTGCTGCT 2396
QY 1812 GAGGCGAGGGATCTACGCCAGGCTGAGGCTTGTGAGGGGTGCGGCTTTCAGCC 1871
Db 2397 GAGGCGAGGGATCTACGCCAGGCTGAGGCTTGTGAGGGGTGCGGCTTTCAGCC 2456
QY 1872 CTCTGGCTTGGCTTTGCTTTCACACGCTG 1899

Db 2457 CTCTGGCTTGGCTTTGCTTTCACACGCTG 2484
RESULT 14
Q0870634 LOCUS 3273 bp DNA linear PAT 13-SEP-2004
DEFINITION Sequence 43 from Patent WO2004073657.
ACCESSION Q0870634
VERSION Q0870634.1 GI:52000145
KEYWORDS .
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS Aizh,N., Gish,K.C., Wilson,K.E. and Zlotnik,A.
TITLE Methods of diagnosis of cancer, composition and methods of
screening for modulators of cancer
JOURNAL Patent: WO 2004073657-A 43 02-SEP-2004;
PROTEIN DESIGN LABS, INC. (US)
FEATURES
source location/Qualifiers
1. .3273
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
ORIGIN
Query Match 92.0%; Score 1748; DB 6; Length 3273;
Best Local Similarity 99.7%; Pred. No. 0;
Matches 1762; Conservative 0; Mismatches 5; Indels 1; Gaps 1;
QY 132 AGCGGCTCCAAACACACACCGCATCTGGAAATTTGGAGCTTCACACCTTCAGCCAGCTGAG 191
Db 1022 ATCGGCTCCAAACACACACCGCATCTGGAAATTTGGAGCT-GACACCTTCAGCCAGCTGAG 1080
QY 192 CTCCTCTCAAAGCCCTGGATCTTAGCTGGAACCGCATCCGGTCCATCCACCTTGAGGCTT 251
Db 1081 CTCCTCTCAAAGCCCTGGATCTTAGCTGGAACCGCATCCGGTCCATCCACCTTGAGGCTT 1140
QY 252 CTCACCCCTGACCTCCCTGCTCAAGCTGAGCCTGACAGACAAACAGCTGACCACTGCTCC 311
Db 1141 CTCACCCCTGACCTCCCTGCTCAAGCTGAGCCTGACAGACAAACAGCTGACCACTGCTCC 1200
QY 312 CTTGGCTGAGCTTTGGGGCTTGTGATCTGAACTCAAAGGAACTTGTCTCTCTCCCA 371
Db 1201 CTTGGCTGAGCTTTGGGGCTTGTGATCTGAACTCAAAGGAACTTGTCTCTCTCCCA 1260
QY 372 GGCCTTCTCAAAGGACAGTTCCTCAAACCTGAGGATCCTTGAGGCTGCTTATGCTTACCA 431
Db 1261 GGCCTTCTCAAAGGACAGTTCCTCAAACCTGAGGATCCTTGAGGCTGCTTATGCTTACCA 1320
QY 432 GTGCTGCTCCTATGAGGATGTGCTGCTTCAAGGCTCTTGGGAGTGGGAGGCTGA 491
Db 1321 GTGCTGCTCCTATGAGGATGTGCTGCTTCAAGGCTCTTCAAGGCTCTTGGGAGTGGGAGGCTGA 1380
QY 492 AGACCTTCACCTTGATGATGAGGATGCTTCAAAGGCGCCCTTGGGCTCTTGGCCAGACA 551
Db 1381 AGACCTTCACCTTGATGATGAGGATGCTTCAAAGGCGCCCTTGGGCTCTTGGCCAGACA 1440
QY 552 AGCAGAGAAACCACTATGACAGGACCTGAGTGAAGTCCAGCTGGAGATGGAGGACTCAAA 611
Db 1441 AGCAGAGAAACCACTATGACAGGACCTGAGTGAAGTCCAGCTGGAGATGGAGGACTCAAA 1500
QY 612 GCCACACCCAGTGTCCAGTGTAGCCCTACTCCAGGCGCCCTTCAAGCCCTGTGTAGTACCT 671
Db 1501 GCCACACCCAGTGTCCAGTGTAGCCCTACTCCAGGCGCCCTTCAAGCCCTGTGTAGTACCT 1560
QY 672 CTTTGAAGCTGGGGCATCCGCTGGCGGTGTGGCCATCGTGTGCTCTCCGTCCTG 731
Db 1561 CTTTGAAGCTGGGGCATCCGCTGGCGGTGTGGCCATCGTGTGCTCTCCGTCCTG 1620
QY 732 CAATGGAAGTGGTGTGCTGCTGACCGTGTTCGCTGGCGGCTGCGCCCTGCGCCCGGTCAA 791

Qy	1872	CTCTGGCTTGGCTTTGCTTCACACGTG	1899
Db	2761	CTCTGGCTTGGCTTTGCTTCACACGTG	2788
RESULT 15			
BC047905			
LOCUS	BC047905	3286 bp	mRNA linear PRI 22-APR-2003
DEFINITION	Homo sapiens leucine-rich repeat-containing G protein-coupled receptor 6, mRNA (cdna clone IMAGE:5212508), partial cds.		
ACCESSION	BC047905		
VERSION	BC047905.1	GI:29126870	
KEYWORDS			
SOURCE	Homo sapiens (human)		
ORGANISM	Homo sapiens		
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.		
REFERENCE	1 (bases 1 to 3286)		
AUTHORS	Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G., Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D., Altshul S.F., Zdobych B., Buetow K.H., Schaefer C.F., Bhat N.K., Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F., Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L., Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E., Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C., Raja S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullay S.J., Bosak S.A., McEwan P.J., McKernan K.J., Malek J.C., Gunaratne P.H., Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W., Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A., Fahey J., Helton E., Ketterman M., Madan A., Rodriguez S., Sanchez A., Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G., Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C., Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S., Krzywinski M.I., Skalska U., Smallos D.E., Schnerch A., Schein J.E., Jones S.J. and Marra M.A.		
TITLE	Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences		
JOURNAL	Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)		
MEDLINE	22388257		
PUBMED	12477932		
REFERENCE	2 (bases 1 to 3286)		
AUTHORS	Strausberg R.		
TITLE	Direct Submission		
JOURNAL	Submitted (03-MAR-2003) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA		
REMARK	NTH-MGC Project URL: http://mgc.nci.nih.gov		
COMMENT	Contact: MGC help desk Email: gcapsb@mail.nih.gov Tissue Procurement: Life Technologies, Inc. cDNA Library Preparation: Life Technologies, Inc. cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL) DNA Sequencing by: Sequencing Group at the Stanford Human Genome Center, Stanford University School of Medicine, Stanford, CA 94305 Web site: http://www-shgc.stanford.edu Contact: (Dickson, Mark) mcd@paxil.stanford.edu R. M., Schmutz, J., Grimwood, J., Rodriguez, A., and Myers, D. M.		
FEATURES	Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov Series: IRAC Plate: 107 Row: D Column: 11. Location/Qualifiers 1. .3286 /organism="Homo sapiens" /mol_type="mRNA" /db_xref="taxon:9606" /clone="IMAGE:5212508" /tissue_type="Blood, adult leukocytes" /clone_lib="NIH MGC 118"		
source			

```

/lab host="DH10B"
/notes="Vector: pCMV-Sport6"
<1..3286
/genes="LGR6"
/notes="synonyms: FLJ14471, VTS20631"
/db_xref="LocusID:59352"
/db_xref="GI:29126871"
<1..2774
/genes="LGR6"
/codon_start=3
/product="LGR6 protein"
/protein_id="AAH47905.1"
/db_xref="GI:29126871"
/db_xref="LocusID:59352"
/db_xref="MIM:606653"
/translation="KGRPTGGMRPRLTLVCVSIISARDLSMNNLTLOPLFLFH
LRFLEURLSNFGLSHLPGFSGLYSLKILMLQNNQNLGGIPAEALWELPSLSLRID
LNLFLSLPERSFGLSLRHLWLDNALTEIPVRALNNLPALQAMTLAARISHIPY
AFQNLTLVLVHLHNNRIQHLGTHSFGLHNLTELDLNYAFOYLPKHLTSLNGAMDIQ
GFHNNITKAIPEKAMGNPLLOTTHFYDNPIQFVGRSAFOYLPKHLTSLNGAMDIQ
FPDLKGTSLLEILTLAGIRLPLSGMCQQLPRLVLELSHNOIEELPSLHRCOKLEE
IGLQHNRIWEGADTFQSLSQALDLSWNAIRSIHPEARSTLSLKLDTDNQLT
LPLAGLGLMHLKUKLNALUSQAFKDSFKPRLILEVPIAYQCPYGMKASFFKASQ
WEABDLHLDDEESKRLGLLARAENHYDQLELQLEMEDSKPSPVQCPPTGPF
KPCYLPFESWIRLAWAI VLLSVLCNGLVLLTVFAGGPVLPVPKPFVGAIAAGNLT
TISGGLLASVDALTFOFSEYGARWETGLGCRATGFLAVLGSEASVLLLTLAAPQCS
TVSVCVAYKSPSLSVRAGLGLALAGLAALPLASVEYGCASVLLCLPYAPPEQO
PAALGPTVALVMNSFELVVGAYIKLYCDLPRGDFEAWDCMWRHVAWLIFADGL
LYCPVAFSLFASMIAGLFPVTPPEAVKSVLLVVLVLPACNLPLLXLLFPHFRDRLRUR
PRAGDSGLPAAVLALESKSDSTQALVAFSDVLILEASEAGRPGLFETGFPSTV
LISCOQAPGRLESHGVEPGRHFNPNQPSMDGELLRLAEGSTPAGGGLSGGGGFPQ
SGLAFASHV"

```

ORIGIN

```

Query Match      92.0%; Score 1748; DB 9; Length 3286;
Best Local Similarity 99.7%; Pred. No. 0;
Matches 1762; Conservative 0; Mismatches 5; Indels 1; Gaps 1;

Qy 132 AGCGGCTCCAGACCAACCGCATCTGGGAATTTGGAGCTNGACACCTTCAGCCAGCTGAG 191
Db 1005 ATCGGCTCCAGACCAACCGCATCTGGGAATTTGGAGCTNGACACCTTCAGCCAGCTGAG 1063

Qy 192 CTCCTCTCAAGCCCTCGATCTTAGCTGGAAGCCATCCGGTCCATCCACCTCAGGCGCTT 251
Db 1064 CTCCTCTCAAGCCCTCGATCTTAGCTGGAAGCCATCCGGTCCATCCACCTCAGGCGCTT 1123

Qy 252 CTCACCCCTGCATCTCCCTGTGCTCAAGCTGAGCTGACAGACCAACCGCTGACCACTGCC 311
Db 1124 CTCACCCCTGCATCTCCCTGTGCTCAAGCTGAGCTGACAGACCAACCGCTGACCACTGCC 1183

Qy 312 CCTGGCTGGACTTGGGGCTTGATGCATCTGAAGCTCAAGGGAACTTGTCTCTCCCA 371
Db 1184 CCTGGCTGGACTTGGGGCTTGATGCATCTGAAGCTCAAGGGAACTTGTCTCTCCCA 1243

Qy 372 GGCCTTCTCCAAAGACAGATTTCCCAAACTCAGGATCTCGAGGTGCTTATGCTTACCA 431
Db 1244 GGCCTTCTCCAAAGACAGATTTCCCAAACTCAGGATCTCGAGGTGCTTATGCTTACCA 1303

Qy 432 GTGCTGCTCCCTATGGGATGTGTGCAGCTTCTTCAAGGCCCTCTGGGAGTGGGAGGCTGA 491
Db 1304 GTGCTGCTCCCTATGGGATGTGTGCAGCTTCTTCAAGGCCCTCTGGGAGTGGGAGGCTGA 1363

Qy 492 AGACCTTCACTTGATGAGGAGTCTTCAAAAGGCCCTTGGGCCCTCTTGGCAGACA 551
Db 1364 AGACCTTCACTTGATGAGGAGTCTTCAAAAGGCCCTTGGGCCCTCTTGGCAGACA 1423

Qy 552 AGCAGAGAACCACTATGACAGGACCTGGATGAGCTCCAGCTGGAGATGGAGGACTCAA 611
Db 1424 AGCAGAGAACCACTATGACAGGACCTGGATGAGCTCCAGCTGGAGATGGAGGACTCAA 1483

Qy 612 GCCACACCCAGTGTCCAGTGTAGCCCTACTCCAGGCCCTTCAAGGCCCTGTGAGTACCT 671
Db 1484 GCCACACCCAGTGTCCAGTGTAGCCCTACTCCAGGCCCTTCAAGGCCCTGTGAGTACCT 1543

```

Qy	672	CTTTGAAAGCTGGGGCATCCGCTGGCCGTGTGGCCCATCGTGTGTCTCTCCGTGCTCTG	731
Db	1544	CTTTGAAAGCTGGGGCATCCGCTGGCCGTGTGGCCCATCGTGTGTCTCTCCGTGCTCTG	1603
Qy	732	CAATGGAATCGTGTCTGCTGACCGGTTCCTGCTGGCGGCGCTGCCGCCCTGCCCCCGGTCAA	791
Db	1604	CAATGGAATCGTGTCTGCTGACCGGTTCCTGCTGGCGGCGCTGCCGCCCTGCCCCCGGTCAA	1663
Qy	792	GTTTGTGGTAGTGTGGATTCGAGGCGCCAAACACCTTGACCTGGGCAATTTCTGTGGCCCTTCT	851
Db	1664	GTTTGTGGTAGTGTGGATTCGAGGCGCCAAACACCTTGACCTGGGCAATTTCTGTGGCCCTTCT	1723
Qy	852	AGCCTCAGTCAGTCCCTGACCTTTTGTGTAGTTCCTGTAGTACGAGAGCCGCTGGGAGAC	911
Db	1724	AGCCTCAGTCAGTCCCTGACCTTTTGTGTAGTTCCTGTAGTACGAGAGCCGCTGGGAGAC	1783
Qy	912	GGGGCTAGGCTGCGGGGCCCACTGGCTTCTGGGAGTACTTTGGGTGCGAGGACATCGGTGCT	971
Db	1784	GGGGCTAGGCTGCGGGGCCCACTGGCTTCTGGGAGTACTTTGGGTGCGAGGACATCGGTGCT	1843
Qy	972	GCTGCTCACTCTGGCCGCGACGTGACGTGACGGTCTCGTCTCTCTGTGTCCGGGCTATGG	1031
Db	1844	GCTGCTCACTCTGGCCGCGACGTGACGTGACGGTCTCGTCTCTCTGTGTCCGGGCTATGG	1903
Qy	1032	GAAGTCCCTCTCTGGGCGAGCGTTTCGAGCAGGGGTCTTAGGCTGCTGGCACATGGCAGG	1091
Db	1904	GAAGTCCCTCTCTGGGCGAGCGTTTCGAGCAGGGGTCTTAGGCTGCTGGCACATGGCAGG	1963
Qy	1092	GCTGGCGCGCGCATGTCCTGCTGCTCAGTGGGAGAAATACGGGGCTTCCCCACATCTGCCT	1151
Db	1964	GCTGGCGCGCGCATGTCCTGCTGCTCAGTGGGAGAAATACGGGGCTTCCCCACATCTGCCT	2023
Qy	1152	GCCTTAGCCGCGACCTGAGGGTTCAGCAGCAGCGCTTGGGCTTACCGTGGCCCTTGTGTAT	1211
Db	2024	GCCTTAGCCGCGACCTGAGGGTTCAGCAGCAGCGCTTGGGCTTACCGTGGCCCTTGTGTAT	2083
Qy	1212	GATGAATCTCTCTGTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	1271
Db	2084	GATGAATCTCTCTGTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	2143
Qy	1272	GCGCGGGGCGACCTTTGAGGCGCTGTTGGGAGCTCGGCCCATGTTGAGGACATGGGCTGGCT	1331
Db	2144	GCGCGGGGCGACCTTTGAGGCGCTGTTGGGAGCTCGGCCCATGTTGAGGACATGGGCTGGCT	2203
Qy	1332	CATCTTTCGAGAGCGGGCTCTCTACTGTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	1391
Db	2204	CATCTTTCGAGAGCGGGCTCTCTACTGTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	2263
Qy	1392	GGGCTCTTCTCCCTGTCAGGCCGCGCTCAAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	1451
Db	2264	GGGCTCTTCTCCCTGTCAGGCCGCGCTCAAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	2323
Qy	1452	GCCTGCTGCTCAACCCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	1511
Db	2324	GCCTGCTGCTCAACCCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	2383
Qy	1512	TCGGCGGCTTTCGGCGCGCGCGAGGAGCTCAGGGGCCCTTAGCCTATGCTGCGGCGCGGGGA	1571
Db	2384	TCGGCGGCTTTCGGCGCGCGCGAGGAGCTCAGGGGCCCTTAGCCTATGCTGCGGCGCGGGGA	2443
Qy	1572	GCTTGGAGAGAGCTCTGTGATTTACCCAGGCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	1631
Db	2444	GCTTGGAGAGAGCTCTGTGATTTACCCAGGCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	2503
Qy	1632	CATTCTGGAGAGCTCTGAAGCTGGGGGCCCTTGGGCTGGAGACCTTATGGCTTCCCTC	1691
Db	2504	CATTCTGGAGAGCTCTGAAGCTGGGGGCCCTTGGGCTGGAGACCTTATGGCTTCCCTC	2563
Qy	1692	AGTGACCTCTATCTCTGTGTCAGCAGCGGGGCCCTTGGGCTGGAGGAGGAGGACCTTGTGT	1751
Db	2564	AGTGACCTCTATCTCTGTGTCAGCAGCGGGGCCCTTGGGCTGGAGGAGGAGGACCTTGTGT	2623

Qy	1752	AGAGCCAGAGGGGAACCACTTTGGGAACCCCAACCTCCATGGATGGAGAACTGCTGCT	1811
Db	2624	AGAGCCAGAGGGGAACCACTTTGGGAACCCCAACCTCCATGGATGGAGAACTGCTGCT	2683
Qy	1812	GAGGGCAGAGGGATCTACGCCAGCAGGTGGAGGCTTGTACAGGGGGTGGCGGCTTTCAGCC	1871
Db	2684	GAGGGCAGAGGGATCTACGCCAGCAGGTGGAGGCTTGTACAGGGGGTGGCGGCTTTCAGCC	2743
Qy	1872	CTCTGGCTTGGCCTTTGCTTTCACACGTG	1899
Db	2744	CTCTGGCTTGGCCTTTGCTTTCACACGTG	2771

Search completed: April 14, 2005, 15:59:16
Job time : 5436.4 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 14, 2005, 10:59:27 ; Search time 669.955 Seconds
(without alignments)
16779.616 Million cell updates/sec

Title: US-10-664-667-6

Perfect score: 1899

Sequence: 1 aatcagctactatagga.....tggcctttgtctcacagtg 1899

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 4390206 seqs, 2959870667 residues

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : N_Geneseq_16Dec04:*

- 1: Geneseqn1980s:*
- 2: Geneseqn1990s:*
- 3: Geneseqn2000s:*
- 4: Geneseqn2001as:*
- 5: Geneseqn2001bs:*
- 6: Geneseqn2002as:*
- 7: Geneseqn2002bs:*
- 8: Geneseqn2003as:*
- 9: Geneseqn2003bs:*
- 10: Geneseqn2003cs:*
- 11: Geneseqn2003ds:*
- 12: Geneseqn2004as:*
- 13: Geneseqn2004bs:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1898	99.9	1899	6	Ai167923 Human LGR
2	1898	99.9	1899	11	Adn02244 Human par
3	1898	99.9	1899	12	Adk19409 ORF of CD
4	1898	99.9	2486	6	Ai167922 Human LGR
5	1898	99.9	2486	11	Adn02242 Human par
6	1898	99.9	2486	12	Adk19407
7	1820.4	95.9	2322	5	Aaf98722 Human lat
8	1811.6	95.4	3451	12	Adq76775 Human LGR
9	1754.4	92.4	2208	6	Ai167925 Human LGR
10	1754.4	92.4	2208	11	Adn02247 Human par
11	1754.4	92.4	2208	12	Adk19412 ORF of pa
12	1754.4	92.4	2711	6	Ai167924 Partial c
13	1754.4	92.4	2711	11	Adn02245 Human par
14	1754.4	92.4	2711	12	Adk19410 Partial c
15	1754.4	92.4	2901	6	Ai167927 Human LGR
16	1754.4	92.4	2901	10	Adc16692 cDNA enco
17	1754.4	92.4	2901	11	Adn02250 Human ful
18	1754.4	92.4	2901	12	Adk19415 ORF of fu
19	1754.4	92.4	3042	10	Adc16698 cDNA enco
20	1754.4	92.4	3325	10	Add89048 Encoding

21	1754.4	92.4	3438	10	ABT31938 Human bre
22	1754.4	92.4	3443	10	Add89090 Encoding
23	1754.4	92.4	3492	6	Ai167926 Human LGR
24	1754.4	92.4	3492	11	Adn02248 Human ful
25	1754.4	92.4	3492	12	Adk19413 Full leng
26	1754.4	92.4	3492	12	Adq15073 Human can
27	1752.8	92.3	2786	10	Adc16710 Human G-p
28	1752.8	92.3	2964	10	Adi21981 Novel hum
29	1752.8	92.3	2988	6	ABZ11841 Human pol
30	1752.8	92.3	2988	12	Adm44359 Novel hum
31	1752.8	92.3	3000	6	Aas98067 Human DNA
32	1752.8	92.3	3042	10	Adc16700 cDNA enco
33	1752.8	92.3	3119	8	ABz42549 Human G p
34	1751.2	92.2	3429	6	AAd37674 Human G-p
35	1749.6	92.1	2487	6	ADH50804 Human G-p
36	1749.6	92.1	2901	10	Adc16695 cDNA with
37	1749.6	92.1	3042	10	Adc16701 cDNA with
38	1748	92.0	2901	10	Adc16694 cDNA enco
39	1748	92.0	2901	10	Adc16696 cDNA enco
40	1748	92.0	3042	10	Adc16702 cDNA with
41	1748	92.0	3273	11	Adn39801 Cancer/an
42	1748	92.0	3273	13	Adr46630 Cancer-as
43	1748	92.0	3306	12	Adq76773 Human LGR
44	1748	92.0	3618	10	Adf70543 Orphan re
45	1746.4	92.0	3349	6	ADH50807 Human G-p

ALIGNMENTS

RESULT 1

AAI67923

ID AAI67923 standard; cDNA; 1899 BP.

XX

AC AAI67923;

XX

DT 13-MAR-2002 (first entry)

XX

DE Human LGR6 polypeptide coding sequence (clone fahr).

XX

KW Large G-protein coupled receptor 6; LGR6; GPCR; immunomodulator;
anti-HIV; cytostatic; neuroprotective; antiparkinsonian;
anticonvulsant; neuroleptic; antidepressant; tranquilizer; antimigraine;
osteopathic; antiarteriosclerotic; vasotropic; hypotensive; cardiant;
antiarrhythmic; anorectic; gene therapy; human; ss.

OS Homo sapiens.

XX

Key Location/Qualifiers

PH 1. .1899

CDS /*tag= a

FT /product= "LGR6 polypeptide"

FT /transl_except= (pos: 169. .171, aa: Xaa)

FT /note= "Xaa = unknown"

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

XX	Example 1; Fig 4; 198pp; English.
PS	The invention provides isolated large G-protein coupled receptor 6 (LGR6)
CC	polypeptides and polynucleotides encoding them. The LGR6 polypeptides and
CC	polynucleotides are useful as targets for developing modulating agents
CC	that regulate a variety of cellular processes, e.g. neural and endocrine
CC	processes, as well as thermogenesis. They are useful for developing novel
CC	diagnostic and therapeutic agents for LGR6 associated disorders such as
CC	weight disorders (anorexia, obesity), eating disorders, acquired
CC	immunodeficiency syndrome (AIDS)-related wasting, cancer-related wasting,
CC	metabolic disorders, central nervous system disorders (e.g. Alzheimer's
CC	disease, Parkinson's disease, Huntington's disease, Gilles de la
CC	Tourette's syndrome, multiple sclerosis, amyotrophic lateral sclerosis,
CC	epilepsy, Jakob-Creutzfeldt disease, depression, schizophrenia,
CC	obsessive-compulsive disorder, anxiety, panic disorder, migraine),
CC	endocrine disorders (hypo and hyperthyroidism, dwarfism, gigantism,
CC	acromegaly), and cardiovascular disorders (atherosclerosis, restenosis,
CC	hypertension, myocardial infarction, arrhythmia). The LGR6 polypeptides,
CC	polynucleotides and antibodies are useful in screening assays, detection
CC	assays (e.g., chromosomal mapping, tissue typing, forensic biology),
CC	predictive medicine (e.g. diagnostic assays, prognostic assays,
CC	monitoring clinical trials and pharmacogenomics), and in methods of
CC	treatment (e.g. therapeutic and prophylactic). The present sequence
CC	represents a human LGR6 polypeptide coding sequence
XX	
SQ	Sequence 1899 BP; 302 A; 624 C; 563 G; 409 T; 0 U; 1 Other;
	Query Match 99.9%; Score 1898; DB 6; Length 1899;
	Best Local Similarity 100.0%; Pred. No. 0;
	Matches 1899; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy	1 AATACGACTCATTATAGGGAAAGCTGTGCGCTGCAGGTACCGTTCGGGAATTCGCCGG 60
Db	1 AATACGACTCATTATAGGGAAAGCTGTGCGCTGCAGGTACCGTTCGGGAATTCGCCGG 60
Qy	61 TCGACCAACGCGTCCGTGGAGCGGACCGAGGTCTGAGCCTGCCGGCTCATCCAGCCTCT 120
Db	61 TCGACCAACGCGTCCGTGGAGCGGACCGAGGTCTGAGCCTGCCGGCTCATCCAGCCTCT 120
Qy	121 CTTGCTGCCCTTAGCGGCTCCAAACAACCGCATCTGGAAAATTGGAGCTNGACACTTC 180
Db	121 CTTGCTGCCCTTAGCGGCTCCAAACAACCGCATCTGGAAAATTGGAGCTNGACACTTC 180
Qy	181 AGCCAGCTAGCTCCCTGCAAGCCCTGGATCTTAGCTGGAACGCCATCCGTTCCATCCAC 240
Db	181 AGCCAGCTAGCTCCCTGCAAGCCCTGGATCTTAGCTGGAACGCCATCCGTTCCATCCAC 240
Qy	241 CCTGAGGCTTCTCCACCCTGCATCCCTGGTCAAGCTTGGACCTTGACAGCAACACGCTG 300
Db	241 CCTGAGGCTTCTCCACCCTGCATCCCTGGTCAAGCTTGGACCTTGACAGCAACACGCTG 300
Qy	301 ACCACATGCCCCCTGCTGGAATTGGGGCTTGATGATCTGAACTCAAAGGAAACCTT 360
Db	301 ACCACATGCCCCCTGCTGGAATTGGGGCTTGATGATCTGAACTCAAAGGAAACCTT 360
Qy	361 GCTCTCTCCAGGCTTCTCCAGGACAGTTTCCAAACTCAGGATCCTGGAGTGCCT 420
Db	361 GCTCTCTCCAGGCTTCTCCAGGACAGTTTCCAAACTCAGGATCCTGGAGTGCCT 420
Qy	421 TATGCTTACAGTGTCTCCCTATGGGATGTGTGCAGCTTCTTCAAGGCTCTGGGCGAG 480
Db	421 TATGCTTACAGTGTCTCCCTATGGGATGTGTGCAGCTTCTTCAAGGCTCTGGGCGAG 480
Qy	481 TGGGAGGCTGAAGACTTCACTTGTATGATGAGGAGTCTTCAAAAGGCCCTTGGGCTC 540
Db	481 TGGGAGGCTGAAGACTTCACTTGTATGATGAGGAGTCTTCAAAAGGCCCTTGGGCTC 540
Qy	541 CTTGGCAGACAGCAGAGAACCATTATGACAGGACCTGGATGAGCTCCAGCTGGAGATG 600
Db	541 CTTGGCAGACAGCAGAGAACCATTATGACAGGACCTGGATGAGCTCCAGCTGGAGATG 600
Qy	601 GAGGACTCAAAGCCAACCCCAGTGTCCAGTGTAGGCTTACTCCAGGCCCCCTTCAAGCCC 660

Qy 781 CCCCGGTCAAGTTTGTGTAGTGCAGTTGCAGGCGCAACACCTTGACTGGCATTTCC 840
Db
Qy 841 TGTGGCTTCTAGGCTCAGTCAATGCTGCTGACCTTTGGTCAAGTCTCTGAGTACGGAGCC 900
Db
Qy 901 CGCTGGGAGACGGGGCTAGGCTGCGGGCCACTGGCTTCTTGGCAGTACTTGGGTGGAG 960
Db
Qy 961 GCATCGGTGCTGCTCACTCTGCGCAGTGCAGTGCAGCTCTCCGTCTCTGCTGTGC 1020
Db
Qy 1021 CGGCGCTATGGAAGTCCCCCTCCCTGGGACGCTTCGAGCAGGGGCTCTAGGCTGGCTG 1080
Db
Qy 1081 GCATCGGTGCTGCTCACTCTGCGCAGTGCAGTGCAGCTCTCCGTCTCTGCTGTGC 1020
Db
Qy 1141 CCATCTGCTGCTGCTCACTCTGCGCAGTGCAGTGCAGCTCTCCGTCTCTGCTGTGC 1020
Db
Qy 1201 GCGGCTATGGAAGTCCCCCTCCCTGGGACGCTTCGAGCAGGGGCTCTAGGCTGGCTG 1080
Db
Qy 1261 GCGGCTATGGAAGTCCCCCTCCCTGGGACGCTTCGAGCAGGGGCTCTAGGCTGGCTG 1080
Db
Qy 1381 GCGGCTATGGAAGTCCCCCTCCCTGGGACGCTTCGAGCAGGGGCTCTAGGCTGGCTG 1080
Db
Qy 1441 GCGGCTATGGAAGTCCCCCTCCCTGGGACGCTTCGAGCAGGGGCTCTAGGCTGGCTG 1080
Db
Qy 1501 GCGGCTATGGAAGTCCCCCTCCCTGGGACGCTTCGAGCAGGGGCTCTAGGCTGGCTG 1080
Db
Qy 1561 GCGGCTATGGAAGTCCCCCTCCCTGGGACGCTTCGAGCAGGGGCTCTAGGCTGGCTG 1080
Db
Qy 1621 GCGGCTATGGAAGTCCCCCTCCCTGGGACGCTTCGAGCAGGGGCTCTAGGCTGGCTG 1080
Db
Qy 1681 GCGGCTATGGAAGTCCCCCTCCCTGGGACGCTTCGAGCAGGGGCTCTAGGCTGGCTG 1080
Db
Qy 1741 GCGGCTATGGAAGTCCCCCTCCCTGGGACGCTTCGAGCAGGGGCTCTAGGCTGGCTG 1080
Db
Qy 1801 GCGGCTATGGAAGTCCCCCTCCCTGGGACGCTTCGAGCAGGGGCTCTAGGCTGGCTG 1080
Db
Qy 1861 GCGGCTATGGAAGTCCCCCTCCCTGGGACGCTTCGAGCAGGGGCTCTAGGCTGGCTG 1080

Db 1861 GGCTTTGAGGCTCTGCTGGCTTTGGCTTTCACACGTG 1899
RESULT 3
ADK19409
ID ADK19409 standard; cDNA; 1899 BP.
XX ADK19409;
AC ADK19409;
XX 17-JUN-2004 (first entry)
XX ORF of cDNA encoding human LGR6 polypeptide #1.
XX G-protein coupled receptor; GPCR; large G-protein coupled receptor 6;
KW LGR6; weight disorder; neural disorder; endocrine disorder;
KW cardiovascular disorder; anorectic; anabolic; immunomodulator;
KW neuroprotective; nootropic; antiparkinsonian; anti-HIV; anticonvulsant;
KW antithyroid; osteopathic; cardiac; vasotonic; hypotensive;
KW antiarteriosclerotic; antiarrhythmic; human; gene; ss.
XX Homo sapiens.
OS
XX
FH Key Location/Qualifiers
CDS 1..1899
FT /*tag= a
FT /partial
FT /product= "LGR6 polypeptide #1"
FT /note= "This sequence lacks both start and stop codons"
FT /transl_except= (pos:169..171,aa:Xaa)
FT /note= "Xaa= Unknown"
XX
DN US2004058377-A1.
XX 25-MAR-2004.
XX 18-SEP-2003; 2003US-00664667.
XX 06-MAY-1999; 99US-0132896P.
PR 08-MAY-2000; 2000US-00566588.
PR 08-MAY-2001; 2001US-00851595.
XX (MILL-) MILLENNIUM PHARM INC.
XX
PI Gu W;
XX
DR WPI; 2004-268789/25.
DR P-PSDB; ADK19408.
XX
XX New large G-protein coupled receptor 6 nucleic acid molecules and
PT polypeptides, useful for diagnosing, preventing or treating diseases
PT associated with aberrant nucleic acid or protein activity, e.g. obesity,
PT anorexia or cachexia.
XX
PS Claim 2; SEQ ID NO 6; 145pp; English.
XX
CC The present invention relates to the isolation of novel members of the G-
CC protein coupled receptor (GPCR) family designated as large G-protein
CC coupled receptor 6 (LGR6), and the polynucleotide sequences encoding
CC them. The invention also discloses LGR6 fusion proteins, antigenic
CC peptides, anti-LGR6 antibodies, recombinant expression vectors, host
CC cells, and non-human transgenic animals in which an LGR6 gene has been
CC introduced or disrupted. The LGR6 polynucleotide sequences, polypeptide
CC sequences, and antibodies are useful for diagnosing, treating or
CC preventing diseases associated with aberrant LGR6 expression or activity,
CC such as weight disorders (e.g. obesity, anorexia nervosa or cachexia),
CC neural disorders (e.g. central nervous system disorder, including
CC Alzheimer's disease, Parkinson's disease, AIDS-related dementia, multiple
CC sclerosis and Huntington's disease), endocrine disorders (e.g.
CC hypothyroidism, hyperthyroidism or acromegaly), or cardiovascular
CC disorders (e.g. ischaemia-reperfusion injury, coronary artery disease,
CC myocardial infarction, arrhythmia, atherosclerosis, hypertension or
CC congestive heart failure). The LGR6 polynucleotide and polypeptide

CC sequences are also useful as targets for developing modulating agents
CC that regulate a variety of cellular processes, e.g. neural and endocrine
CC processes and thermogenesis. They can be used in screening assays (e.g.
CC chromosome mapping, tissue typing or in forensic biology), or in
CC predictive medicine (e.g. diagnostic assays, prognostic assays,
CC monitoring clinical trials or pharmacogenomics). The present sequence
CC encodes a human LOR6 polypeptide.

XX	SQ	Sequence 1899 BP; 302 A; 624 C; 563 G; 409 T; 0 U; 1 Other;	
		Query Match 99.9%; Score 1898; DB 12; Length 1899;	
		Best Local Similarity 100.0%; Pred. No. 0;	
		Matches 1899; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	
Qy	1	AATAGACTCATTATAGGAAAGCTGGTACGGCTCAGGTACCGTCCGGAAATCCCGGG	60
Db	1	AATAGACTCATTATAGGAAAGCTGGTACGGCTCAGGTACCGTCCGGAAATCCCGGG	60
Qy	61	TCGACCCACGGTCCGTGGAGCGGAGCCAGGGTCTGAGCTCCGGCTCATCCAGCTCT	120
Db	61	TCGACCCACGGTCCGTGGAGCGGAGCCAGGGTCTGAGCTCCGGCTCATCCAGCTCT	120
Qy	121	CTTGCTGCCCTAGCGGCTCCAAACAACCGCATCTGGAAATTCGAGCTNGACACCTTC	180
Db	121	CTTGCTGCCCTAGCGGCTCCAAACAACCGCATCTGGAAATTCGAGCTNGACACCTTC	180
Qy	181	AGCCAGCTGAGCTCCCTGCAAGCCCTGGATCTTAGCTGAAAGCCATCCGTTCCATCCAC	240
Db	181	AGCCAGCTGAGCTCCCTGCAAGCCCTGGATCTTAGCTGAAAGCCATCCGTTCCATCCAC	240
Qy	241	CCTGAGGCTTCTCCACCTGTCATCTCCCTGGTCAAGCTGAGCTGACAGCAACAGCTG	300
Db	241	CCTGAGGCTTCTCCACCTGTCATCTCCCTGGTCAAGCTGAGCTGACAGCAACAGCTG	300
Qy	301	ACCACATGCCCTGGCTGGACTTGGGGCTTGATGCATCTGAAGCTCAAGGGAACTT	360
Db	301	ACCACATGCCCTGGCTGGACTTGGGGCTTGATGCATCTGAAGCTCAAGGGAACTT	360
Qy	361	GCTCTCTCCAGGCTTCTCCAAAGACAGTTTCCAAAACTGAGGATCTTGAGGTGCT	420
Db	361	GCTCTCTCCAGGCTTCTCCAAAGACAGTTTCCAAAACTGAGGATCTTGAGGTGCT	420
Qy	421	TATGCTTACAGTGTGCTCCCTATGGGATGTGTGTCAGCTTCTCAAGGCTCTGGGAG	480
Db	421	TATGCTTACAGTGTGCTCCCTATGGGATGTGTGTCAGCTTCTCAAGGCTCTGGGAG	480
Qy	481	TGGGAGCTGAAGCTTTCACCTTGATGAGGAGTCTTCAAAAGGCCCTTGGGCTC	540
Db	481	TGGGAGCTGAAGCTTTCACCTTGATGAGGAGTCTTCAAAAGGCCCTTGGGCTC	540
Qy	541	CTTGCCAGACAGAGAACCACTATGACAGGACCTGGATGAGCTCCAGCTGGAGATG	600
Db	541	CTTGCCAGACAGAGAACCACTATGACAGGACCTGGATGAGCTCCAGCTGGAGATG	600
Qy	601	GAGGACTCAAAAGCCACACCCAGTGTCCAGTGTAGCCCTTACTCCAGGCCCCCTTCAAGCC	660
Db	601	GAGGACTCAAAAGCCACACCCAGTGTCCAGTGTAGCCCTTACTCCAGGCCCCCTTCAAGCC	660
Qy	661	TGTGAGTACCTTTTGAAGCTGGGGCATCCGCTTGGCGTGGGGCCATCGTGTGCTC	720
Db	661	TGTGAGTACCTTTTGAAGCTGGGGCATCCGCTTGGCGTGGGGCCATCGTGTGCTC	720
Qy	721	TCCGTGCTCTGCAATGGACTGTGCTGACCGTGTTCGCTGGGGCTTCCGCCCTG	780
Db	721	TCCGTGCTCTGCAATGGACTGTGCTGACCGTGTTCGCTGGGGCTTCCGCCCTG	780
Qy	781	CCCCCGGTCAAGTTTGT	840
Db	781	CCCCCGGTCAAGTTTGT	840
Qy	841	TGTGCTCTTACAGCTCAGTCCATGCTGACCTTTGGTCACTTCTCTGAGTACGAGCC	900
Db	841	TGTGCTCTTACAGCTCAGTCCATGCTGACCTTTGGTCACTTCTCTGAGTACGAGCC	900

RESULT 4
AAI67922
ID AAI67922 standard; cDNA; 2486 BP.

Qy	901	CGCTGGAGACGGGGCTAGGCTGCGGGCCACTGCTTCTTGCGCAGTACTTGGGTGGAG	960
Db	901	CGCTGGAGACGGGGCTAGGCTGCGGGCCACTGCTTCTTGCGCAGTACTTGGGTGGAG	960
Qy	961	GCATCGGTGCTGCTGCTCACTCTGGCGCAGTGCAGTGCAGGCTCTCGTCTCTGTC	1020
Db	961	GCATCGGTGCTGCTGCTCACTCTGGCGCAGTGCAGTGCAGGCTCTCGTCTCTGTC	1020
Qy	1021	CGGGCTATGGAAAGTCCCTCTGCGCAGCGTTTCAGCAGGGGCTCTTAGGCTGCCCTG	1080
Db	1021	CGGGCTATGGAAAGTCCCTCTGCGCAGCGTTTCAGCAGGGGCTCTTAGGCTGCCCTG	1080
Qy	1081	GCATCGCAGGGCTGGCGCGCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	1140
Db	1081	GCATCGCAGGGCTGGCGCGCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	1140
Qy	1141	CCACTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	1200
Db	1141	CCACTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	1200
Qy	1201	GCTCTGGTGTATGAACTCTTCTGTTTCTGTTTCTGTTTCTGTTTCTGTTTCTGTTT	1260
Db	1201	GCTCTGGTGTATGAACTCTTCTGTTTCTGTTTCTGTTTCTGTTTCTGTTTCTGTTT	1260
Qy	1261	TACTGTGACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	1320
Db	1261	TACTGTGACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	1320
Qy	1321	GTGGCTGGCTCATCTTTCGAGAGCGGCTCTTCTTCTGCTGCTGCTGCTGCTGCTGCTG	1380
Db	1321	GTGGCTGGCTCATCTTTCGAGAGCGGCTCTTCTTCTGCTGCTGCTGCTGCTGCTGCTG	1380
Qy	1381	GCCTCATGCTGGGCTCTTCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	1440
Db	1381	GCCTCATGCTGGGCTCTTCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	1440
Qy	1441	GTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	1500
Db	1441	GTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	1500
Qy	1501	CGGATGACCTTTCGGGCTTTCGGGCTTTCGGGCTTTCGGGCTTTCGGGCTTTCGGGCT	1560
Db	1501	CGGATGACCTTTCGGGCTTTCGGGCTTTCGGGCTTTCGGGCTTTCGGGCTTTCGGGCT	1560
Qy	1561	CGGCGGGGAGCTGGAGAGAGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	1620
Db	1561	CGGCGGGGAGCTGGAGAGAGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	1620
Qy	1621	GATGTGATCTCATTTCTGAAAGCTTCTGAAAGCTTCTGAAAGCTTCTGAAAGCTTCTG	1680
Db	1621	GATGTGATCTCATTTCTGAAAGCTTCTGAAAGCTTCTGAAAGCTTCTGAAAGCTTCTG	1680
Qy	1681	GGCTTTCCTCAGTGAACCTCATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	1740
Db	1681	GGCTTTCCTCAGTGAACCTCATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	1740
Qy	1741	AGCCATTTGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG	1800
Db	1741	AGCCATTTGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG	1800
Qy	1801	GAACTGCTGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG	1860
Db	1801	GAACTGCTGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG	1860
Qy	1861	GGCTTTTACAGCCCTCTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGG	1899
Db	1861	GGCTTTTACAGCCCTCTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGG	1899

Db 1082 GCATGGCAGGGCTGGCGCCGCACTGCCCTTGGCCCTCAGTGGGAGATATACGGGGCCCTCC 1141
Qy 1141 CCACCTCTGCTGCTCCTAGCGCCACCTCAGGGTCAAGCAGCAGCCCTGGGCTTACCGGTG 1200
Db 1142 CCACCTCTGCTGCTCCTAGCGCCACCTCAGGGTCAAGCAGCAGCCCTGGGCTTACCGGTG 1201
Qy 1201 GCCCTGGTATGATGAATCT 1260
Db 1202 GCCCTGGTATGATGAATCT 1261
Qy 1261 TACTGTGACCTGCGCGGGGACATTTGAGGCGGCTGAGGCGGCTGAGGCGCATGCTGAGGCGAC 1320
Db 1262 TACTGTGACCTGCGCGGGGACATTTGAGGCGGCTGAGGCGGCTGAGGCGCATGCTGAGGCGAC 1321
Qy 1321 GTGGCTGGCTCATCTTCCGACAGCGGCTCTCTACTGTCCGCTGGCCCTTCTCTCAGCTTC 1380
Db 1322 GTGGCTGGCTCATCTTCCGACAGCGGCTCTCTACTGTCCGCTGGCCCTTCTCTCAGCTTC 1381
Qy 1381 GCCTCATGCTGGGCTCTTCCCTGTACGCCCGGAGGCGGTCAAGTCTGTCTCTCTCTCTCTCT 1440
Db 1382 GCCTCATGCTGGGCTCTTCCCTGTACGCCCGGAGGCGGTCAAGTCTGTCTCTCTCTCTCTCT 1441
Qy 1441 GTGCTGCCCTGCCCTGCTGCTCAACCACTGCTACCTGCTCTTCAACCCCGACCTTC 1500
Db 1442 GTGCTGCCCTGCCCTGCTGCTCAACCACTGCTACCTGCTCTTCAACCCCGACCTTC 1501
Qy 1501 CGGATGACCTTGGCGGCTTCCGCGCCCGCGAGGCGACTCAGGCGCCCTAGCCTATGCT 1560
Db 1502 CGGATGACCTTGGCGGCTTCCGCGCCCGCGAGGCGACTCAGGCGCCCTAGCCTATGCT 1561
Qy 1561 GCGGCGGGGAGCTGGAAGAGCTCTCTGTGATTTACCCAGGCGCTGTAGCCTTCTCT 1620
Db 1562 GCGGCGGGGAGCTGGAAGAGCTCTCTGTGATTTACCCAGGCGCTGTAGCCTTCTCT 1621
Qy 1621 GATGTGATCTCATTTCTGGAAGCTTCTGAGCTGGCGGCGCCCTTGGGCTGGAGACCTAT 1680
Db 1622 GATGTGATCTCATTTCTGGAAGCTTCTGAGCTGGCGGCGCCCTTGGGCTGGAGACCTAT 1681
Qy 1681 GCTTCCCTCAGTCACTCTATCTCTCTGTCAGCAGCAGGCGGCGCCCTGAGGCGGCT 1740
Db 1682 GCTTCCCTCAGTCACTCTATCTCTCTGTCAGCAGCAGGCGGCGCCCTGAGGCGGCT 1741
Qy 1741 AGCCATTGTAGAGCCAGAGGGGAACCACTTTGGGAACCCCAACCTTCCATGGATGGA 1800
Db 1742 AGCCATTGTAGAGCCAGAGGGGAACCACTTTGGGAACCCCAACCTTCCATGGATGGA 1801
Qy 1801 GAACTGCTGTAGGCGCAGAGGATCTACGCCAGCAGGTGAGGCTTGTACGGGGTGGC 1860
Db 1802 GAACTGCTGTAGGCGCAGAGGATCTACGCCAGCAGGTGAGGCTTGTACGGGGTGGC 1861
Qy 1861 GCTTTACGCCCTCTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCT 1899
Db 1862 GCTTTACGCCCTCTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCT 1900

RESULT 5
ADN02242

ID ADN02242 standard; cDNA; 2486 BP.
XX AC ADN02242;
XX DT 17-JUN-2004 (first entry)
XX DE Human partial large G protein-coupled receptor, LGR6, cDNA #1.
XX KW Human; ss; large G protein-coupled receptor 6; LGR6; heart failure;
KW ischaemic heart disease; myocardial infarction; hypertension;
KW pericarditis; atherosclerosis; hepatic failure; viral hepatitis; cancer;
KW intracerebral haemorrhage; brain abscess; Alzheimer's disease;
KW Pick disease; Parkinson's disease; Huntington's disease;
KW amyotrophic lateral sclerosis; inflammatory bowel disease;
KW Crohn's disease; benign prostatic hypertrophy;
KW systemic lupus erythematosus.

XX OS Homo sapiens.
XX FH Key Location/Qualifiers
FT CDS 2..1903
FT /*tag=a
FT /partial
FT /note= "No start codon shown"
FT /product= "Partial LGR6"
XX US2003166047-A1.
XX PD 04-SEP-2003.
XX PF 08-MAY-2001; 2001US-00851595.
XX PR 06-MAY-1999; 99US-0132896P.
XX PR 08-MAY-2000; 2000US-00566588.
XX PA (MILL-) MILLENNIUM PHARM INC.
XX PI Gu W;
XX WI WPI; 2003-898067/82.
XX DR P-PSDB; ADN02243.
XX PT New isolated LGR6 nucleic acids and polypeptides, useful for diagnosing and treating LGR6-mediated disorders, such as myocardial infarction, hypertension, atherosclerosis, viral hepatitis, cancer and/or Alzheimer's disease.
XX PS Claim 1; SEQ ID NO 4; 145pp; English.
XX CC The invention relates to an isolated LGR6 (large G protein-coupled receptor) nucleic acid molecule comprising the mouse and human cDNA sequences appearing as ADN02239, ADN02241, ADN02242, ADN02244, ADN02245, ADN02247, ADN02248 and ADN02250. Also included are a host cell or a non-human mammalian host cell containing an LGR6 nucleic acid (and producing LGR6), an isolated LGR6 polypeptide appearing as ADN02240, ADN02243, ADN02246 and ADN02249, an antibody which selectively binds to LGR6, a method for detecting the presence of LGR6 in a sample, a kit (comprising a compound which selectively binds to LGR6, or hybridises to LGR6 nucleic acid, and instructions for use), a method for detecting LGR6 nucleic acid in a sample, a method for identifying a compound which binds to LGR6, a method for modulating the activity of LGR6 and a method for identifying a compound that modulates the activity of LGR6. The methods and compositions of the present invention are useful for diagnosing and treating large G-protein coupled receptor 6 (LGR6)-mediated or -related disorders, such as heart failure, ischaemic heart disease, myocardial infarction, hypertension, pericarditis, atherosclerosis, hepatic failure, viral hepatitis, cancer, intracerebral haemorrhage, brain abscess, Alzheimer's disease, Pick disease, Parkinson's disease, Huntington's disease, amyotrophic lateral sclerosis, inflammatory bowel disease, Crohn's disease, benign prostatic hypertrophy and systemic lupus erythematosus. The present sequence is a partial human cDNA for LGR6.
XX SQ Sequence 2486 BP; 455 A; 773 C; 692 G; 565 T; 0 U; 1 Other;
Query Match 99.9%; Score 1898; DB 11; Length 2486;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1899; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 AATACGACTCACTATAGGAAAGCTGCTACGCCCTGACGGTACCGGTCCGGAAATTCCTCCGG 60
Db 2 AATACGACTCACTATAGGAAAGCTGCTACGCCCTGACGGTACCGGTCCGGAAATTCCTCCGG 61
Qy 61 TCGACCCACGCGTCCGTGGAGCGGAGCGAGGCTCTGAGCCTTCGCGGCTCATCCAGCCTCT 120
Db 62 TCGACCCACGCGTCCGTGGAGCGGAGCGAGGCTCTGAGCCTTCGCGGCTCATCCAGCCTCT 121
Qy 121 CTTGCTGCCCTAGCGGCTTCCAAACACACCGCATCTGGGAAATTTGGAGCTNGACACCTTC 180
Db 122 CTTGCTGCCCTAGCGGCTTCCAAACACACCGCATCTGGGAAATTTGGAGCTNGACACCTTC 181

181 AGCCAGCTGAGCTCCCTGCRAGCCCTGGATCTTAGCTGGAACGCCATCCGTCATCCAC 240
 182 AGCCAGCTGAGCTCCCTGCAAGCCCTGGATCTTAGCTGGAACGCCATCCGTCATCCAC 241
 241 CCTGAGGCTTCTCCACCCCTGACCTCCCTGGTCAAGCTGGAACCTGACAGAACACAGCTG 300
 242 CCTGAGGCTTCTCCACCCCTGACCTCCCTGGTCAAGCTGGAACCTGACAGAACACAGCTG 301
 301 ACCACATGCCCCCTGGCTGGAATTGGGGCTTGGATGCAATCTGAAGCTCAAAAGGAACTTT 360
 302 ACCACATGCCCCCTGGCTGGAATTGGGGCTTGGATGCAATCTGAAGCTCAAAAGGAACTTT 361
 361 GCTCTCTCCAGGCTTCTCCAAAGGACAGTTTCCAAAACCTGAGATCCTGGAGGTGCCT 420
 362 GCTCTCTCCAGGCTTCTCCAAAGGACAGTTTCCAAAACCTGAGATCCTGGAGGTGCCT 421
 421 TATGCTACCAAGTGTCTCCCTATGGGATGTGTGCCAGCTTCTTCAAGGCCCTCTGGGCAG 480
 422 TATGCTACCAAGTGTCTCCCTATGGGATGTGTGCCAGCTTCTTCAAGGCCCTCTGGGCAG 481
 481 TGGAGGCTGAAGACCTTCACTTGTATGAGGAGTCTTCAAAAAGGCCCTCTGGGCCTC 540
 482 TGGAGGCTGAAGACCTTCACTTGTATGAGGAGTCTTCAAAAAGGCCCTCTGGGCCTC 541
 541 CTTGCCAGACAGCAGAGAACCTATGACAGACCTGGATGAGCTCCAGCTGGAGATG 600
 542 CTTGCCAGACAGCAGAGAACCTATGACAGACCTGGATGAGCTCCAGCTGGAGATG 601
 601 GAGGACTCAAAAGCACACCCAGTGTCCAGTGTAGCCCTACTCCAGGCCCTTCAAGGCC 660
 602 GAGGACTCAAAAGCACACCCAGTGTCCAGTGTAGCCCTACTCCAGGCCCTTCAAGGCC 661
 661 TGTGAGTACCTCTTTGAAAGCTGGGGCATCCGCTGCGGTGGGCGATCGTGTGCTC 720
 662 TGTGAGTACCTCTTTGAAAGCTGGGGCATCCGCTGCGGTGGGCGATCGTGTGCTC 721
 721 TCCGTGCTCTGCAATGAGCTGGTCTGCTGACCGTGTGCTGGGGGCTGCCCCCTG 780
 722 TCCGTGCTCTGCAATGAGCTGGTCTGCTGACCGTGTGCTGGGGGCTGCCCCCTG 781
 781 CCCCCTGTCAGTTGTGTAGTGGATGCGATTGCGAGCGCCCAACACTTGAAGCTTCC 840
 782 CCCCCTGTCAGTTGTGTAGTGGATGCGATTGCGAGCGCCCAACACTTGAAGCTTCC 841
 841 TGTGGCTTCTAGGCTCAGTGCATGCGCTGACCTTTGGTCAAGTCTCTGAGTACGAGCC 900
 842 TGTGGCTTCTAGGCTCAGTGCATGCGCTGACCTTTGGTCAAGTCTCTGAGTACGAGCC 901
 901 CGCTGGGAGACGGGGCTAGGCTGCGGGCCACTGGCTTCTGGCAGTACTTGGGTGCGAG 960
 902 CGCTGGGAGACGGGGCTAGGCTGCGGGCCACTGGCTTCTGGCAGTACTTGGGTGCGAG 961
 961 GCATCGGTGCTGCTGCTCACTCTGGCGCAGTGCAGTGCAGCGTCTCGTCTCCTGTGTC 1020
 962 GCATCGGTGCTGCTGCTCACTCTGGCGCAGTGCAGTGCAGCGTCTCGTCTCCTGTGTC 1021
 1021 CGGCGCTATGGAAGTCCCTCCCTGGCAGCGTTCGAGCAGGGCTCTAGGCTGCGCTG 1080
 1022 CGGCGCTATGGAAGTCCCTCCCTGGCAGCGTTCGAGCAGGGCTCTAGGCTGCGCTG 1081
 1081 GCATCGGAGGGCTGGGCGCGCATGCTGCCCTGGCCCTCAGTGGGAGATACGGGGCTCC 1140
 1082 GCATCGGAGGGCTGGGCGCGCATGCTGCCCTGGCCCTCAGTGGGAGATACGGGGCTCC 1141
 1141 CCACCTGCTGCTGCTAGCGGCCACTGAGGGTCAAGCAGAGCCCTGGGCTTCAACCGTG 1200
 1142 CCACCTGCTGCTGCTAGCGGCCACTGAGGGTCAAGCAGAGCCCTGGGCTTCAACCGTG 1201
 1201 GCGCTGTGATGATGAATCTCTTCTGTTTCTGTTGCGGTGCTCATCAAACTG 1260
 1202 GCGCTGTGATGATGAATCTCTTCTGTTTCTGTTGCGGTGCTCATCAAACTG 1261

QY 1261 TACTGTGACCTGCGCGGGCGGACTTTGAGGCGGTGTGGACTGCGCCATGGTGGAGCAC 1320
 Db 1262 TACTGTGACCTGCGCGGGCGGACTTTGAGGCGGTGTGGACTGCGCCATGGTGGAGCAC 1321
 QY 1321 GTGGCTTGGCTCATCTTTCGAGACAGGGCTCCTCTACTGTCCCGTGGCTTCTCAGCTTC 1380
 Db 1322 GTGGCTTGGCTCATCTTTCGAGACAGGGCTCCTCTACTGTCCCGTGGCTTCTCAGCTTC 1381
 QY 1381 GCCTTCATGCTGGGCTCTTCCCTGTCACGCCGAGGCCGCTCAAGTCTGTCTGCTGCTG 1440
 Db 1382 GCCTTCATGCTGGGCTCTTCCCTGTCACGCCGAGGCCGCTCAAGTCTGTCTGCTGCTG 1441
 QY 1441 GTGCTGCCCTGCTGCTGCTCAACCCACTGCTGTACCTGTCTTCAACCCCACTTC 1500
 Db 1442 GTGCTGCCCTGCTGCTGCTCAACCCACTGCTGTACCTGTCTTCAACCCCACTTC 1501
 QY 1501 CGGGATGACCTTTCGGGGCTTTCGGGCCCGCGCAGGGGACTTCAGGGGCCCTTAGCCTATGCT 1560
 Db 1502 CGGGATGACCTTTCGGGGCTTTCGGGCCCGCGCAGGGGACTTCAGGGGCCCTTAGCCTATGCT 1561
 QY 1561 GCGGCGGGGAGCTGGAGAGAGCTCCTGTGATTCTACCCAGGCCCTGTAGCTTCTCT 1620
 Db 1562 GCGGCGGGGAGCTGGAGAGAGCTCCTGTGATTCTACCCAGGCCCTGTAGCTTCTCT 1621
 QY 1621 GATGTGGATCTCATTTCTGGAAGCTTCTGAAGCTGGGCGGCCCTTGGGCTGGAGACCTAT 1680
 Db 1622 GATGTGGATCTCATTTCTGGAAGCTTCTGAAGCTGGGCGGCCCTTGGGCTGGAGACCTAT 1681
 QY 1681 GGCTTCCCTCAGTACCTCATCTCTGTGACGACCCAGGGGCCCTTGGCTGGAGGGC 1740
 Db 1682 GGCTTCCCTCAGTACCTCATCTCTGTGACGACCCAGGGGCCCTTGGCTGGAGGGC 1741
 QY 1741 AGCCATGTGTAGACCCAGAGGGGACCTTTGGGAACCCCAACCTCCATGGATGGA 1800
 Db 1742 AGCCATGTGTAGACCCAGAGGGGACCTTTGGGAACCCCAACCTCCATGGATGGA 1801
 QY 1801 GAACTGTCTGAGGGCAGAGGATCTACGCCACAGAGTGGAGCTTGTGAGGGGTGGC 1860
 Db 1802 GAACTGTCTGAGGGCAGAGGATCTACGCCACAGAGTGGAGCTTGTGAGGGGTGGC 1861
 QY 1861 GGCTTTACGCGCTTGGCTTGGCTTGTCTTCAACGTC 1899
 Db 1862 GGCTTTACGCGCTTGGCTTGGCTTGTCTTCAACGTC 1900

RESULT 6
 ADK19407
 ID ADK19407 standard; cDNA; 2486 BP.
 XX
 AC ADK19407;
 XX
 DX 17-JUN-2004 (first entry)
 XX
 cDNA encoding human LGR6 polypeptide #1.
 DE
 XX
 KW G-protein coupled receptor; GPCR; large G-protein coupled receptor 6;
 KW LGR6; weight disorder; neural disorder; endocrine disorder;
 KW cardiovascular disorder; anorectic; anabolic; immunomodulator;
 KW neuroprotective; nootropic; antiparkinsonian; anti-HIV; anticonvulsant;
 KW antithyroid; osteopathic; cardiac; vasotropic; hypotensive;
 KW antiarteriosclerotic; antiarrhythmic; human; gene; ss.
 XX
 OS Homo sapiens.
 XX
 Key Location/Qualifiers
 CDS 2..1903
 FT /*tag= a
 FT /partial
 FT /product= "LGR6 polypeptide #1"
 FT /note= "This sequence lacks a start codon"
 FT /transl_except= (pos:170..172,aa:Xaa)
 FT /note= "Xaa= Unknown"
 XX

PN US2004058377-A1.
XX PD 25-MAR-2004.
XX PF 18-SEP-2003; 2003US-00664667.
XX PR 06-MAY-1999; 99US-0132896P.
XX PR 08-MAY-2000; 2000US-00566588.
XX PR 08-MAY-2001; 2001US-00851595.
XX PA (MILL-) MILLENNIUM PHARM INC.
XX PI Gu W;
XX DR WPI; 2004-268789/25.
XX DR P-PSDB; ADK19408.
XX PR New large G-protein coupled receptor 6 nucleic acid molecules and
PT polypeptides, useful for diagnosing, preventing or treating diseases
PT associated with aberrant nucleic acid or protein activity, e.g. obesity,
PT anorexia or cachexia.
XX Claim 2; SEQ ID NO 4; 145pp; English.
XX The present invention relates to the isolation of novel members of the G-
CC protein coupled receptor (GPCR) family designated as large G-protein
CC coupled receptor 6 (LGR6), and the polynucleotide sequences encoding
CC them. The invention also discloses LGR6 fusion proteins, antigenic
CC peptides, anti-LGR6 antibodies, recombinant expression vectors, host
CC cells, and non-human transgenic animals in which an LGR6 gene has been
CC introduced or disrupted. The LGR6 polynucleotide sequences, polypeptide
CC sequences, and antibodies are useful for diagnosing, treating or
CC preventing diseases associated with aberrant LGR6 expression or activity,
CC such as weight disorders (e.g. obesity, anorexia nervosa or cachexia),
CC neural disorders (e.g. central nervous system disorder, including
CC Alzheimer's disease, Parkinson's disease, AIDS-related dementia, multiple
CC sclerosis and Huntington's disease), endocrine disorders (e.g.
CC hypothyroidism, hyperthyroidism or acromegaly), or cardiovascular
CC disorders (e.g. leukaemia-reperfusion injury, coronary artery disease,
CC myocardial infarction, arrhythmia, atherosclerosis, hypertension or
CC congestive heart failure). The LGR6 polynucleotide and polypeptide
CC sequences are also useful as targets for developing modulating agents
CC that regulate a variety of cellular processes, e.g. neural and endocrine
CC processes and thermogenesis. They can be used in screening assays (e.g.
CC chromosome mapping, tissue typing or in forensic biology), or in
CC predictive medicine (e.g. diagnostic assays, prognostic assays,
CC monitoring clinical trials or pharmacogenomics). The present sequence
CC encodes a human LGR6 polypeptide.
XX Sequence 2486 BP; 455 A; 773 C; 692 G; 565 T; 0 U; 1 Other;
Query Match 99.9%; Score 1898; DB 12; Length 2486;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1899; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 AATAGACTCACTATAGGAAAGCTGGTACGCTCGAGTACCGTCCGGAAATCCCGGG 60
DB 2 AATAGACTCACTATAGGAAAGCTGGTACGCTCGAGTACCGTCCGGAAATCCCGGG 61
QY 61 TCGACCCACGCTCCGTGAGCGGAGCCAGGGTCTGAGCTGCGGCTCATCCAGCTCT 120
DB 62 TCGACCCACGCTCCGTGAGCGGAGCCAGGGTCTGAGCTGCGGCTCATCCAGCTCT 121
QY 121 CTTGTGCTCCCTAGCGGCTCCAAACACCCGATCTGGGAATTCGAGCTGACACCTTC 180
DB 122 CTTGTGCTCCCTAGCGGCTCCAAACACCCGATCTGGGAATTCGAGCTGACACCTTC 181
QY 181 AGCCAGCTGAGCTCCCTGCAAGCCCTGGATCTTAGCTGGAAGCCATCCGTTCCAC 240
DB 182 AGCCAGCTGAGCTCCCTGCAAGCCCTGGATCTTAGCTGGAAGCCATCCGTTCCAC 241
QY 241 CTTGAGGCTTCTCCACCCCTGCATCTCCCTGCTCAAGCTGGAGCTGACAGAACCCAGCTG 300

DB 242 CTTGAGGCTTCTCCACCCCTGCATCTCCCTGCTCAAGCTGGACCTGACAGACACACGCTG 301
QY 301 ACCACACTGCCCCCTGGCTGGACTTTGGGGCTTTGATGATCTGAAGCTCAAAAGGGAACCTT 360
DB 302 ACCACACTGCCCCCTGGCTGGACTTTGGGGCTTTGATGATCTGAAGCTCAAAAGGGAACCTT 361
QY 361 GCTCTCTCCAGGCTTCTCCAGGACAGTTTCCAAAGCTGAGGATCCTGGAGTGCTT 420
DB 362 GCTCTCTCCAGGCTTCTCCAGGACAGTTTCCAAAGCTGAGGATCCTGGAGTGCTT 421
QY 421 TATGCTTACCACTGCTGCTCCCTATGGGATGTCGACAGCTTCTTCAAGSCCTCTGGGGCAG 480
DB 422 TATGCTTACCACTGCTGCTCCCTATGGGATGTCGACAGCTTCTTCAAGSCCTCTGGGGCAG 481
QY 481 TGGGAGGCTGAAGACCTTCACTTATGATGAGGAGTCTTCAAAAAGGCCCTCTGGGGCTC 540
DB 482 TGGGAGGCTGAAGACCTTCACTTATGATGAGGAGTCTTCAAAAAGGCCCTCTGGGGCTC 541
QY 541 CTTGCCAGACAAAGCAGAGAACCACTATGACAGGACCTGGATGAGCTCCAGCTGGAGATG 600
DB 542 CTTGCCAGACAAAGCAGAGAACCACTATGACAGGACCTGGATGAGCTCCAGCTGGAGATG 601
QY 601 GAGGACTCAAGGCCACACCCCACTGTCAGTGTAGCCCTACTCCAGGSCCCCTTCAAGGCC 660
DB 602 GAGGACTCAAGGCCACACCCCACTGTCAGTGTAGCCCTACTCCAGGSCCCCTTCAAGGCC 661
QY 661 TGTGAGTACCTCTTTGAAAGCTGGGGCATCCGCTGGCCGTGTGGGCCATCGTGTGCTC 720
DB 662 TGTGAGTACCTCTTTGAAAGCTGGGGCATCCGCTGGCCGTGTGGGCCATCGTGTGCTC 721
QY 721 TCCGTGCTCTCAATGAGACTGGTCTGTGACCTGTCACCTGTTGCTGGCGGCCCTGCCCTC 780
DB 722 TCCGTGCTCTCAATGAGACTGGTCTGTGACCTGTCACCTGTTGCTGGCGGCCCTGCCCTC 781
QY 781 CCCCCGCTCAAGTTTGTGAGTGGCATTTCAGGCGCCACACACTTTCAGCTGGGCATTTCC 840
DB 782 CCCCCGCTCAAGTTTGTGAGTGGCATTTCAGGCGCCACACACTTTCAGCTGGGCATTTCC 841
QY 841 TGTGGCTTCTAGGCTCAGTGCATGCTGCTGACCTTTGGTTCAGTTCCTCTGAGTACGAGGCC 900
DB 842 TGTGGCTTCTAGGCTCAGTGCATGCTGCTGACCTTTGGTTCAGTTCCTCTGAGTACGAGGCC 901
QY 901 CGCTGGGAGACGGGGCTAGGCTGCGGGCCACTGGCTTCTGGCAGTACTTTGGGTGCGAG 960
DB 902 CGCTGGGAGACGGGGCTAGGCTGCGGGCCACTGGCTTCTGGCAGTACTTTGGGTGCGAG 961
QY 961 GCATCGGTGCTGCTGCTCACTCTGGCGCAGTGCAGTGCAGGCTCTCGTCTCTGCTGCTC 1020
DB 962 GCATCGGTGCTGCTGCTCACTCTGGCGCAGTGCAGTGCAGGCTCTCGTCTCTGCTGCTC 1021
QY 1021 CGGGCTATGGGAAGTCCCTCCCTGGGCGAGCGTTTCAGCAGGGGCTCTAGGCTGCGCTG 1080
DB 1022 CGGGCTATGGGAAGTCCCTCCCTGGGCGAGCGTTTCAGCAGGGGCTCTAGGCTGCGCTG 1081
QY 1081 GCATCGGAGGGCTGGCGCGCCACTGCCCCTGGCCTCAGTGGGAGAAATACGGGGCTTCC 1140
DB 1082 GCATCGGAGGGCTGGCGCGCCACTGCCCCTGGCCTCAGTGGGAGAAATACGGGGCTTCC 1141
QY 1141 CCATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1200
DB 1142 CCATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1201
QY 1201 GCCCTGGTGTATGAACTCTCTTCTGTTTCTGGTCTGGGCGGCTGCTGCTGCTGCTGCTGCTG 1260
DB 1202 GCCCTGGTGTATGAACTCTCTTCTGTTTCTGGTCTGGGCGGCTGCTGCTGCTGCTGCTGCTG 1261
QY 1261 TACTGTGACCTGCGCGGGCGAGCTTTGAGGCCGTGTGGGACTGCGCCATGGTGGAGCAC 1320
DB 1262 TACTGTGACCTGCGCGGGCGAGCTTTGAGGCCGTGTGGGACTGCGCCATGGTGGAGCAC 1321
QY 1321 GTGGCTGGCTCATCTTCGAGAGAGGGCTCTCTACTGTCTCCGCTGGCCCTTCTCTGAGCTTC 1380
DB 1322 GTGGCTGGCTCATCTTCGAGAGAGGGCTCTCTACTGTCTCCGCTGGCCCTTCTCTGAGCTTC 1381

QY 1381 GCCTCATGCTGGGCTCTTCCCTGTCACGCCGAGGCGTCAAGTCTGTCTCTGTG 1440
 Db 1382 GCCTCATGCTGGGCTCTTCCCTGTCACGCCGAGGCGTCAAGTCTGTCTCTGTG 1441
 QY 1441 GTGTGCCCCCTGCTGCTGCTCAACCCACTGTCTTCAACCCCACTTC 1500
 Db 1442 GTGTGCCCCCTGCTGCTGCTCAACCCACTGTCTTCAACCCCACTTC 1501
 QY 1501 CGGGATGACCTTGGCGGCTTGGCCCCGCGAGGGACTCAGGCCCTAGCCTATGCT 1560
 Db 1502 CGGGATGACCTTGGCGGCTTGGCCCCGCGAGGGACTCAGGCCCTAGCCTATGCT 1561
 QY 1561 GGGCGCGGAGCTGAGAGAGCTCTGTGATTCTACCCAGGCCCTGTAGCTTCTCT 1620
 Db 1562 GGGCGCGGAGCTGAGAGAGCTCTGTGATTCTACCCAGGCCCTGTAGCTTCTCT 1621
 QY 1621 GATGTGATCTCATTTCTGAAAGCTTCTGAAGCTGGCGGCCCTTGGCTGGAGACCTAT 1680
 Db 1622 GATGTGATCTCATTTCTGAAAGCTTCTGAAGCTGGCGGCCCTTGGCTGGAGACCTAT 1681
 QY 1681 GCTTTCCTCTAGTACCTCTATCTCTGTGAGCAGCCAGGGCCCCCAGGCTGGAGGCG 1740
 Db 1682 GCTTTCCTCTAGTACCTCTATCTCTGTGAGCAGCCAGGGCCCCCAGGCTGGAGGCG 1741
 QY 1741 AGCCATTGTGTAGAGCAGAGGGGAACCACTTTTGGGAACCCCAACCTTCCATGGATGA 1800
 Db 1742 AGCCATTGTGTAGAGCAGAGGGGAACCACTTTTGGGAACCCCAACCTTCCATGGATGA 1801
 QY 1801 GAACTGCTGTGAGGCGAGAGGATCTACGCCAGCAGGTGGAGGCTTGTCAAGGGGTTGC 1860
 Db 1802 GAACTGCTGTGAGGCGAGAGGATCTACGCCAGCAGGTGGAGGCTTGTCAAGGGGTTGC 1861
 QY 1861 GCTTTAGCCCTCTGGCTTGGCTTTGCTTGCATCAGCTG 1899
 Db 1862 GCTTTAGCCCTCTGGCTTGGCTTTGCTTGCATCAGCTG 1900

RESULT 7

AAF98722
 ID AAF98722 standard; DNA; 2322 BP.
 XX
 AC AAF98722;
 XX
 DT 02-JUL-2001 (first entry)
 XX
 DE Human late stage ovarian tumour polynucleotide marker 37.
 XX
 KW Human; ovarian cancer; identification; detection; characterisation;
 KW tumour; kinase; marker; cytostatic; antisense gene therapy; ds.
 XX
 OS Homo sapiens.
 XX
 XX WO200118542-A2.
 XX
 PD 15-MAR-2001.
 XX
 PF 01-SEP-2000; 2000WO-US024199.
 XX
 PR 03-SEP-1999; 99US-0152547P.
 PR 16-MAR-2000; 2000US-0190347P.
 PR 21-MAR-2000; 2000US-0191321P.
 PR 31-MAY-2000; 2000US-0208382P.
 PR 20-JUL-2000; 2000US-0022046P.
 XX
 PA (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.
 XX
 PI Lee J, Thompso P, Lillie J;
 XX
 DR WPI; 2001-211428/21.
 XX
 PT Detection, assessment, prevention and therapy of ovarian cancer,
 PT comprises detecting changes in the expression of a variety of markers.

XX Claim 1; Page 1188-1189; 1198pp; English.
 PS The present invention describes a method for assessing whether a patient
 CC is afflicted with ovarian cancer by comparing: (1) the expression of a
 CC marker (1) (see AAF98594 to AAF98730), in a patient sample; and (2) the
 CC normal level of expression of (1) in a control non-ovarian cancer sample,
 CC where a significant difference between the level of expression in (a) and
 CC (b) is an indication that the patient is afflicted with ovarian cancer.
 CC (1) have cytostatic activities and can be used in antisense gene therapy.
 CC The method, compositions and kits from the present invention can be used
 CC for: (1) assessing and treating ovarian cancer; (2) making isolated
 CC hybridoma, which produces an antibody useful for ovarian cancer
 CC assessment; and (3) inhibiting ovarian cancer in a patient. AAF98573 to
 CC AAF98593 represent human kinase marker primers and probes which are used
 CC in the exemplification of the present invention
 XX
 SQ Sequence 2322 BP; 407 A; 733 C; 652 G; 530 T; 0 U; 0 Other;
 Query Match 95.9%; Score 1820.4; DB 5; Length 2322;
 Best Local Similarity 99.9%; Pred. No. 0;
 Matches 1832; Conservative 0; Mismatches 1; Indels 1; Gaps 1;
 QY 66 CCACGCTGCTGAGGCGGAGCCAGGCTCTGAGCTGCGGCTCATCCAGCTCTCTTGC 125
 Db 1 CCACGCTGCTGAGGCGGAGCCAGGCTCTGAGCTGCGGCTCATCCAGCTCTCTTGC 60
 QY 126 TGCCTAGCGGCTCTCAACACACCGCATCTGGGAAATTTGAGCTNACACCTTCAGCCA 185
 Db 61 TGCCTAGCGGCTCTCAACACACCGCATCTGGGAAATTTGAGCTNACACCTTCAGCCA 119
 QY 186 GCTGAGCTCCCTGCAAGCCCTGGATCTTAGCTGAAGCCATCCGCTCCATCCACCCTGA 245
 Db 120 GCTGAGCTCCCTGCAAGCCCTGGATCTTAGCTGAAGCCATCCGCTCCATCCACCCTGA 179
 QY 246 GGCCTTCTCCACCTGCACTCCCTGCTCAAGCTGGACCTGACAGAACACAGCTGACAC 305
 Db 180 GGCCTTCTCCACCTGCACTCCCTGCTCAAGCTGGACCTGACAGAACACAGCTGACAC 239
 QY 306 ACTGCCCCCTGGCTGGACTTGGGGCTTGTGATGATCTGAAGCTCAAGGAAACCTTGTCT 365
 Db 240 ACTGCCCCCTGGCTGGACTTGGGGCTTGTGATGATCTGAAGCTCAAGGAAACCTTGTCT 299
 QY 366 CTCCAGGCTTCTCCAGGACAGTTTCCAAACCTGAGGATCTCGAGGTGCTTATGC 425
 Db 300 CTCCAGGCTTCTCCAGGACAGTTTCCAAACCTGAGGATCTCGAGGTGCTTATGC 359
 QY 426 CTACAGTGTCTCCCTATGGGATGTGTGCGAGCTTCTTCAAGGCTCTGGGCAAGTGGGA 485
 Db 360 CTACAGTGTCTCCCTATGGGATGTGTGCGAGCTTCTTCAAGGCTCTGGGCAAGTGGGA 419
 QY 486 GGTGAAGACCTTCACTTGTATGATGAGGATCTTCAAAAGGCGCTTGGGCTCTTGC 545
 Db 420 GGTGAAGACCTTCACTTGTATGATGAGGATCTTCAAAAGGCGCTTGGGCTCTTGC 479
 QY 546 CAGACAGCAGAGAACCACTATGACAGGACCTGGATGAGCTCCAGCTGGAGATGGAGGA 605
 Db 480 CAGACAGCAGAGAACCACTATGACAGGACCTGGATGAGCTCCAGCTGGAGATGGAGGA 539
 QY 606 CTAAAGCCACACCCCAAGTGTCCAGTGTAGCCCTTACTCCAGGCGCTTCAAGCCCTTGA 665
 Db 540 CTAAAGCCACACCCCAAGTGTCCAGTGTAGCCCTTACTCCAGGCGCTTCAAGCCCTTGA 599
 QY 666 GTACCTCTTTGAAAGCTGGGGCATCCGCTGCGGCTGTGGGCGCATTCGTGTCTCCGT 725
 Db 600 GTACCTCTTTGAAAGCTGGGGCATCCGCTGCGGCTGTGGGCGCATTCGTGTCTCCGT 659
 QY 726 GCTCTCAATGGACTGTGCTGCTGACCGTGTGCTGGCGGCTCCGCCCTGCCCCC 785
 Db 660 GCTCTCAATGGACTGTGCTGCTGACCGTGTGCTGGCGGCTCCGCCCTGCCCCC 719
 QY 786 GGTCAAGTCTTGTGTAGGTGCGATTGACGGCGCAACACCTTGTGACTGGCATTTCTGTGG 845

Db 720 GGTCAAGTTTGTGGTAGTGGGATTGTCAGGCGCCCAACACCTTGAAGTGGCAATTTCTCTGTGG 779
Qy 846 CTTCTTAGCCTCAGTCGATGCCCTTGACCTTTGGTCAGTTCTCTGAGTAGCGAGCCCGCTG 905
Db 780 CTTCTTAGCCTCAGTCGATGCCCTTGACCTTTGGTCAGTTCTCTGAGTAGCGAGCCCGCTG 839
Qy 906 GGAGACGGGGCTAGGCTGCGGGCCCACTGGGCTTCTGGCAGTACCTTGGGTGGAGGCATC 965
Db 840 GGAGACGGGGCTAGGCTGCGGGCCCACTGGGCTTCTGGCAGTACCTTGGGTGGAGGCATC 899
Qy 966 GGTGCTGTCTCACTCTGGCGCAGTGCAGTGCAGCTCTCCGTCCTCTGTGTGTCGGGC 1025
Db 900 GGTGCTGTCTCACTCTGGCGCAGTGCAGTGCAGCTCTCCGTCCTCTGTGTGTCGGGC 959
Qy 1026 CTATGGGAAGTCCCTCCCTCGGGCAGCGTTTCGACAGGGGTCTTAGGCTGCCTGGCACT 1085
Db 960 CTATGGGAAGTCCCTCCCTCGGGCAGCGTTTCGACAGGGGTCTTAGGCTGCCTGGCACT 1019
Qy 1086 GGCAGGGCTGGCGCGCACTGCCCTGGCCTCAGTGGGAGAAATACGGGGCTTCCCACT 1145
Db 1020 GGCAGGGCTGGCGCGCACTGCCCTGGCCTCAGTGGGAGAAATACGGGGCTTCCCACT 1079
Qy 1146 CTGCTGCGCTACGCGCCACCTGAGGCTCAGCCAGCAGCCCTGGGCTTCAAGTGGCCCT 1205
Db 1080 CTGCTGCGCTACGCGCCACCTGAGGCTCAGCCAGCAGCCCTGGGCTTCAAGTGGCCCT 1139
Qy 1206 GGTGATGATGAATCTCTTCTGTTCTGCTGGTGGCGGTGCTACATCAAACTGACTG 1265
Db 1140 GGTGATGATGAATCTCTTCTGTTCTGCTGGTGGCGGTGCTACATCAAACTGACTG 1199
Qy 1266 TGACCTGCGCGGGCGCACTTTGAGCCGTGTGGCACTGCGCCATGTTGAGCACGTGGC 1325
Db 1200 TGACCTGCGCGGGCGCACTTTGAGCCGTGTGGCACTGCGCCATGTTGAGCACGTGGC 1259
Qy 1326 CTGGCTCATCTCCGACAGGGCTCTCTACTGTCCGTGGCTTCCCTCAGCTTGGCTC 1385
Db 1260 CTGGCTCATCTCCGACAGGGCTCTCTACTGTCCGTGGCTTCCCTCAGCTTGGCTC 1319
Qy 1386 CATGCTGGGCTCTTCCCTGTACGCCCGAGGCGCTCAAGTCTGCTGCTGGTGGTCT 1445
Db 1320 CATGCTGGGCTCTTCCCTGTACGCCCGAGGCGCTCAAGTCTGCTGCTGGTGGTCT 1379
Qy 1446 GCCCTGCTGCTGCTCAACCCCACTGCTGTACCTGCTCTTCAACCCCACTTCCGGGA 1505
Db 1380 GCCCTGCTGCTGCTCAACCCCACTGCTGTACCTGCTCTTCAACCCCACTTCCGGGA 1439
Qy 1506 TGACCTTGGGGCTTCCGGCCCGGCGAGGGACTCAGGGCCCTAGGCTATGCTGGGC 1565
Db 1440 TGACCTTGGGGCTTCCGGCCCGGCGAGGGACTCAGGGCCCTAGGCTATGCTGGGC 1499
Qy 1566 CGGGAGCTGGAGAGAGCTCTGTTGATTCTACCCAGGCGCTGGTAGCTTCTCTGATGT 1625
Db 1500 CGGGAGCTGGAGAGAGCTCTGTTGATTCTACCCAGGCGCTGGTAGCTTCTCTGATGT 1559
Qy 1626 GGATCTCATTTGGAAGCTTCTGAAGCTGGGCGGCCCTCGGGCTGGAGACCTATGGCTT 1685
Db 1560 GGATCTCATTTGGAAGCTTCTGAAGCTGGGCGGCCCTCGGGCTGGAGACCTATGGCTT 1619
Qy 1686 CCCCTCAGTACCTCATCTCTGTACAGCAGCGGGCCCGCCAGGCTGGAGGGAGCGCA 1745
Db 1620 CCCCTCAGTACCTCATCTCTGTACAGCAGCGGGCCCGCCAGGCTGGAGGGAGCGCA 1679
Qy 1746 TTGTGTAGCCAGAGGGGACCACTTTGGGAAACCCCAACCTCCATGGATGGAGACT 1805
Db 1680 TTGTGTAGCCAGAGGGGACCACTTTGGGAAACCCCAACCTCCATGGATGGAGACT 1739
Qy 1806 GCTGCTAGGGCAGAGGGATCTACGCCAGAGGTGGAGGCTTGTTCAGGGGGTGGCGGCTT 1865
Db 1740 GCTGCTAGGGCAGAGGGATCTACGCCAGAGGTGGAGGCTTGTTCAGGGGGTGGCGGCTT 1799
Qy 1866 TCAGCCCTCTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTT 1899
Db 1800 TCAGCCCTCTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTT 1833

RESULT 8

ADQ76775
ID ADQ76775 standard; DNA; 3451 BP.

XX AC ADQ76775;

XX DT 07-OCT-2004 (first entry)

XX Human LGR6 splice variant LGR6.2 nucleotide sequence.

XX Human; leucine-rich repeat-containing G-protein coupled receptor;
receptor; LGR6.2; cytostatic; immunomodulator; muscular-gen.;
gynaecological; antiinfertility; endocrine-gen.; anabolic; hypertensive;
vulnary; dermatological; gene therapy; gene; ds.

OS Homo sapiens.

XX Key Location/Qualifiers
CDS 47..1093
/tag= b

FT sig_peptide /product= "Human LGR6.2"

FT 47..103 /tag= a

FT mat_peptide 104..1090 /tag= c

PN WO2004058818-A2.

XX 15-JUL-2004.

XX 23-DEC-2003; 2003WO-EP051091.

XX 26-DEC-2002; 2002US-0436001P.

XX (ISTF) ARS APPLIED RES SYSTEMS HOLDING NV.

XX Muda M;

XX WPI; 2004-534126/51.

XX Novel isolated leucine-rich repeat-containing G-protein coupled receptor-
6 splicing variant 1 and 2 polypeptides, useful for treating or
preventing endometriosis, female or male infertility, Addison's disease,
Cushing's disease, cancers.

XX Claim 1; SEQ ID NO 4; 85pp; English.

XX The present nucleotide sequence encodes human LGR6.2, a novel splice
variant of leucine-rich repeat-containing G-protein coupled receptor 6
(LGR6-SV). It was initially identified as an expressed sequence tag
clone. The LGR6.2 coding sequence consists of the same exons present in
another splice variant, LGR6.1 ADQ76773, plus an extended exon that
introduces a stop codon resulting in truncation of the coding sequence,
removing the transmembrane domains and the cytoplasmic C-terminal domain
of LGR6.1. Thus, LGR6.2 is a secreted, probably soluble, N-terminal
extracellular domain version of LGR6.1 and probably functions as an
antagonist of the LGR6.1 signalling pathway. Tissue distribution studies
indicate a high level of expression in ovary and pituitary tissues. The N
-terminal extracellular domain of LGR6.2 can be used as an antagonist of
the LGR6 signalling pathway. The invention provides LGR6.1 and LGR6.2
polypeptides and nucleic acids, as well as selective binding agents. Also
vectors, host cells and methods for producing LGR6-SV polypeptides. Also
provided are pharmaceutical compositions and methods for the diagnosis,
treatment, amelioration and/or prevention of diseases, disorders and
conditions associated with LGR6-SV polypeptides. LGR6-SV polypeptides and
nucleic acids can be used for diagnosing or treating cancers such as
hepatocellular carcinoma, adenomatous polypsis, malignant transformation
of epithelial cells, colorectal cancer, breast cancer, ovarian cancer,
etc.; cachexia and muscular dystrophy; miscarriage, endometriosis,
uterine cancer, and female infertility; Cushing's disease and Addison's
disease; male infertility and testicular cancer; leukaemia; and diseases

CC and conditions affecting the kidney. They may also be useful for treating
CC diseases and conditions involving cell proliferation and differentiation,
CC such as tissue damage and degeneration, ageing, wound healing, cancer,
CC hyperplasia and hypertrophy.

XX	SQ	Sequence 3451 BP; 728 A; 1063 C; 920 G; 740 T; 0 U; 0 Other;	
		Query Match 95.4%; Score 1811.6; DB 12; Length 3451;	
		Best Local Similarity 99.7%; Pred. No. 0;	
		Matches 1825; Conservative 0; Mismatches 4; Indels 1; Gaps 1;	
Qy	70	GGGTCCGTGGAGCGGAGCAGGGTCTGAGCTGCGCGCTCATCCAGCCCTCTTGTCTGCC	129
Db	1066	GGGGCATGGAGCGAGCAGGGTCTGAGCTGCGCGCTCATCCAGCCCTCTTGTCTGCC	1125
Qy	130	CTAGCGGCTCACAACACCGCATCTGGGAAATTGGAGCTNGACACTTCAGCCAGCTG	189
Db	1126	CTAGCGGCTCACAACACCGCATCTGGGAAATTGGAGCT-GACACCTTCAGCCAGCTG	1184
Qy	190	AGCTCCCTGCAAGCCCTGGATCTTAGCTGGACGCATCCGGTCCATCCACCTGAGGCC	249
Db	1185	AGCTCCCTGCAAGCCCTGGATCTTAGCTGGAAACGCCATCCGGTCCATCCACCTGAGGCC	1244
Qy	250	TTCTCCACCTGCACTCCCTGGTCAAGCTGACCTGACAGACAACACAGCTGACACACTG	309
Db	1245	TTCTCCACCTGCACTCCCTGGTCAAGCTGACCTGACAGACAACACAGCTGACACACTG	1304
Qy	310	CCCTGGCTGACATTTGGGGCTTGATGATCTGAAGCTCAAGGAACTTGTCTCTCC	369
Db	1305	CCCTGGCTGACATTTGGGGCTTGATGATCTGAAGCTCAAGGAACTTGTCTCTCC	1364
Qy	370	CAGGCTTCTCCMAGGACAGTTTCCCAAACTGAGGATCCTGGAGTGCCTTATGCCCTAC	429
Db	1365	CAGGCTTCTCCMAGGACAGTTTCCCAAACTGAGGATCCTGGAGTGCCTTATGCCCTAC	1424
Qy	430	CAGTCTGTCCCTATGGGATGTGCCAGCTTCTTCAAGGCTCTTGGGAGTGGAGGCT	489
Db	1425	CAGTCTGTCCCTATGGGATGTGCCAGCTTCTTCAAGGCTCTTGGGAGTGGAGGCT	1484
Qy	490	GAAGACCTTCACTTGTATGATGAGAGTCTTCAAAAAGGCCCTCTGGGCTCTTGGCCAGA	549
Db	1485	GAAGACCTTCACTTGTATGATGAGAGTCTTCAAAAAGGCCCTCTGGGCTCTTGGCCAGA	1544
Qy	550	CAAGCAGAGACCACTATGACAGACCTGAGTCCAGTCCAGCTGGAGTGGAGACTCA	609
Db	1545	CAAGCAGAGAACCACTATGACAGACCTGAGTCCAGTCCAGCTGGAGTGGAGACTCA	1604
Qy	610	AAGCCACACCCAGTGTCCAGTGTAGCCCTACTCCAGGCCCTTCAAGCCCTGTGAGTAC	669
Db	1605	AAGCCACACCCAGTGTCCAGTGTAGCCCTACTCCAGGCCCTTCAAGCCCTGTGAGTAC	1664
Qy	670	CTCTTTGAAAGCTGGGGCATCCGCTGCGCGTGGCCATCGTGTGCTCTCCGTGCTC	729
Db	1665	CTCTTTGAAAGCTGGGGCATCCGCTGCGCGTGGCCATCGTGTGCTCTCCGTGCTC	1724
Qy	730	TGCAATGACATGTGCTGCTGACCGTGTTCGCTGGCGGGCTTGCCTCCCTCCCGGTC	789
Db	1725	TGCAATGACATGTGCTGCTGACCGTGTTCGCTGGCGGGCTTGCCTCCCTCCCGGTC	1784
Qy	790	AAGTTTGTGTAGGTGGATTTGAGGCGCCAAACCTTGTGCTGGCATTTCTGTGGCCTT	849
Db	1785	AAGTTTGTGTAGGTGGATTTGAGGCGCCAAACCTTGTGCTGGCATTTCTGTGGCCTT	1844
Qy	850	CTAGCCTCAGTCGATGCCCTGACCTTTGGTTCAGTTCTCTGAGTACGAGCCCGCTGGAG	909
Db	1845	CTAGCCTCAGTCGATGCCCTGACCTTTGGTTCAGTTCTCTGAGTACGAGCCCGCTGGAG	1904
Qy	910	ACGGGGCTAGGCTGCGGGCCACTGGCTTCTGGCAGTACTTGGTTCGAGGCTACCGTG	969
Db	1905	ACGGGGCTAGGCTGCGGGCCACTGGCTTCTGGCAGTACTTGGTTCGAGGCTACCGTG	1964
Qy	970	CTGCTGTCTCACTTGGCGCGCAGTGCAGTGCAGCGCTTCCGTCTCTGTGTCCGGCCCTAT	1029

RESULT 9

AAI67925

ID AAI67925 standard; cDNA; 2208 BP.

XX AAI67925;

XX 13-MAR-2002 (first entry)

XX Human LGR6 polypeptide coding sequence.

Db	1965	CTGCTGTCTCACTCTGGCGCGCAGTGCAGTGCAGCGTCTCCGTCTCTGTGTCCGGCCCTAT	2024
Qy	1030	GGGAAGTCCCTCTCCCTGGGGCAGGTTTCGAGCAGGGGTCTTAGGCTGCTGGCACTGGCA	1089
Db	2025	GGGAAGTCCCTCTCCCTGGGGCAGGTTTCGAGCAGGGGTCTTAGGCTGCTGGCACTGGCA	2084
Qy	1090	GGGTGCGCGCGCAGTGCCTGGCTCAGTGGGAGAAACACGGGGCTCCGCCACTCTGC	1149
Db	2085	GGGTGCGCGCGCAGTGCCTGGCTCAGTGGGAGAAACACGGGGCTCCGCCACTCTGC	2144
Qy	1150	CTGCCCTACGCGCACCTGAGGGTTCAGCCAGCAGCCCTGGGCTTTCACCGTGGCCCTGGTG	1209
Db	2145	CTGCCCTACGCGCACCTGAGGGTTCAGCCAGCAGCCCTGGGCTTTCACCGTGGCCCTGGTG	2204
Qy	1210	ATGATGAATCTCTTCTGTTCCTGGTCTGGCGGGTCTACATCAAACTGTACTGTGAC	1269
Db	2205	ATGATGAATCTCTTCTGTTCCTGGTCTGGCGGGTCTACATCAAACTGTACTGTGAC	2264
Qy	1270	CTGCCGCGGGCGGACTTTGAGGCGCTGCGGACCTGCGCCATGGTGGGACGCTGGCCCTGG	1329
Db	2265	CTGCCGCGGGCGGACTTTGAGGCGCTGCGGACCTGCGCCATGGTGGGACGCTGGCCCTGG	2324
Qy	1330	CTCATCTTCGAGACGCGGCTCTCTACTGTCCCGTGGCCTTTCCTCAGCTTCGCCCTCCATG	1389
Db	2325	CTCATCTTCGAGACGCGGCTCTCTACTGTCCCGTGGCCTTTCCTCAGCTTCGCCCTCCATG	2384
Qy	1390	CTGGGCTCTTCCCTGTTCAGCGCGGAGCGTCAAGTCTGTCTCTGTCTGTGGTGTCTGCC	1449
Db	2385	CTGGGCTCTTCCCTGTTCAGCGCGGAGCGTCAAGTCTGTCTCTGTCTGTGGTGTCTGCC	2444
Qy	1450	CTGCTCTCCCTCAACCCACTGCTGTACTGTCTTCAACCCCTTTCAGGCTTCCCGGATGAC	1509
Db	2445	CTGCTCTCCCTCAACCCACTGCTGTACTGTCTTCAACCCCTTTCAGGCTTCCCGGATGAC	2504
Qy	1510	CTTCCGCGGCTTCGCGCCCGCAGGGGACTCAGGGGCCCTTAGCCTATGTCTGCGGCCGG	1569
Db	2505	CTTCCGCGGCTTCGCGCCCGCAGGGGACTCAGGGGCCCTTAGCCTATGTCTGCGGCCGG	2564
Qy	1570	GAGCTGAGAGAGCTCTCTGTGATTTACCGAGGCCCTGTAGGCTTCTCTGATGTGGAT	1629
Db	2565	GAGCTGAGAGAGAGCTCTCTGTGATTTACCGAGGCCCTGTAGGCTTCTCTGATGTGGAT	2624
Qy	1630	CTCAATCTGGAAGCTTCTGAAGCTGGCGCGCCCTGGGCTGGAGACTTATGGCTTCCCC	1689
Db	2625	CTCAATCTGGAAGCTTCTGAAGCTGGCGCGCCCTGGGCTGGAGACTTATGGCTTCCCC	2684
Qy	1690	TCAGTGACCTCTCTCTGTCTGACGACGAGGGGCCCGCCAGGCTGGAGGCGAGCATTGT	1749
Db	2685	TCAGTGACCTCTCTCTGTCTGACGACGAGGGGCCCGCCAGGCTGGAGGCGAGCATTGT	2744
Qy	1750	GTAGAGCAGAGGGGAACCACTTTGGGAACCCCGAACCCCTCCATGGATGGAGAACTGCTG	1809
Db	2745	GTAGAGCAGAGGGGAACCACTTTGGGAACCCCGAACCCCTCCATGGATGGAGAACTGCTG	2804
Qy	1810	CTGAGGSCACAGGAGTCTACGACGAGGTGGAGGCTTGTACGGGGTGGCGGCTTTCAG	1869
Db	2805	CTGAGGSCACAGGAGTCTACGACGAGGTGGAGGCTTGTACGGGGTGGCGGCTTTCAG	2864
Qy	1870	CCCTCTGGCTTGGCTTTGCTTTCACACGCTG	1899
Db	2865	CCCTCTGGCTTGGCTTTGCTTTCACACGCTG	2894

Large G-protein coupled receptor 6; LGR6; GPCR; immunomodulator; anti-HIV; cytostatic; nootropic; neuroprotective; antiparkinsonian; anticonvulsant; neuroleptic; antidepressant; tranquilizer; antimigraine; osteopathic; antiarteriosclerotic; vasotropic; hypotensive; cardiant; antiarrhythmic; anorectic; gene therapy; human; ss.

Homo sapiens.

Key Location/Qualifiers
CDS 1..2208
/*tag= a
/product= "LGR6 polypeptide"

WO200185768-A2.
15-NOV-2001.
08-MAY-2001; 2001WO-US015002.
08-MAY-2000; 2000US-00566588.
(MILL-) MILLENNIUM PHARM INC.
Gu W;
WPI; 2002-055584/07.
P-PSDB; AAG66140.

Novel isolated large G-protein coupled receptor 6 polypeptide, useful for diagnosing and treating weight disorder, metabolic disorder, central nervous system disorder, endocrine disorder and cardiovascular disorder.

Claim 2; Fig 8; 198pp; English.

The invention provides isolated large G-protein coupled receptor 6 (LGR6) polypeptides and polynucleotides encoding them. The LGR6 polypeptides and polynucleotides are useful as targets for developing modulating agents that regulate a variety of cellular processes, e.g. neural and endocrine processes, as well as thermogenesis. They are useful for developing novel diagnostic and therapeutic agents for LGR6 associated disorders such as weight disorders (anorexia, obesity), eating disorders, acquired immunodeficiency syndrome (AIDS)-related wasting, cancer-related wasting, metabolic disorders, central nervous system disorders (e.g. Alzheimer's disease, Parkinson's disease, Huntington's disease, Gilles de la Tourette's syndrome, multiple sclerosis, amyotrophic lateral sclerosis, epilepsy, Jakob-Creutzfeldt disease, depression, schizophrenia, obsessive-compulsive disorder, anxiety, panic disorder, migraine), endocrine disorders (hypo and hyperthyroidism, dwarfism, gigantism, acromegaly), and cardiovascular disorders (atherosclerosis, restenosis, hypertension, myocardial infarction, arrhythmia). The LGR6 polypeptides, polynucleotides and antibodies are useful in screening assays, detection assays (e.g. chromosomal mapping, tissue typing, forensic biology), predictive medicine (e.g. diagnostic assays, prognostic assays, monitoring clinical trials and pharmacogenomics), and in methods of treatment (e.g. therapeutic and prophylactic). The present sequence represents a human LGR6 polypeptide coding sequence

Sequence 2208 BP; 396 A; 710 C; 631 G; 471 T; 0 U; 0 Other;

Query Match 92.4%; Score 1754.4; DB 6; Length 2208;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1766; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

132 AGCGGCTCCACACACCGCATCGGAAATGGAGCTNGACACCTTCAGCCAGCTGAG 191
133 |
442 ATCGGCTCCACACACCGCATCGGAAATGGAGCT-GACACCTTCAGCCAGCTGAG 500
134 |
192 CTCCTGCAAGCCCTGGATCTTAGCTGGAAGCCATCCGGTCCATCCACCTCAGGCGCTT 251
193 |
501 CTCCTGCAAGCCCTGGATCTTAGCTGGAAGCCATCCGGTCCATCCACCTCAGGCGCTT 560
502 |
252 CTCACCTGCACTCCCTGGTCAAGCTGGACCTGACAGACCAACCAAGCTGACCACTGCC 311
253 |

Db 561 CTCACCCCTGCATCCCTCGTCAAGCTGGACCTGACAGACCAACAGCTGACCACTGCC 620
Qy 312 CCTGGCTGGACTTGGGGCTTTGATGCATCTCAAGCTCAAGGGGAACCTTGTCTCTCCCA 371
Db 621 CCTGGCTGGACTTGGGGCTTTGATGCATCTCAAGCTCAAGGGGAACCTTGTCTCTCCCA 680
Qy 372 GGCCTTCTCCAAAGACAGTTTCCCAAAACTGAGGATCTCGAGGTGCTTATGCTATACA 431
Db 681 GGCCTTCTCCAAAGACAGTTTCCCAAAACTGAGGATCTCGAGGTGCTTATGCTATACA 740
Qy 432 GTGTGTCTCTATGGGATGTGCAGCTTCTCAAGGCTCTGGGCAAGTGGAGGCTGA 491
Db 741 GTGTGTCTCTATGGGATGTGCAGCTTCTCAAGGCTCTGGGCAAGTGGAGGCTGA 800
Qy 492 AGACCTTTCACCTTGATGATGAGGAGTCTTCAAAAGAGCCCTTGGGCTCTCTTGCAGACA 551
Db 801 AGACCTTTCACCTTGATGATGAGGAGTCTTCAAAAGAGCCCTTGGGCTCTCTTGCAGACA 860
Qy 552 AGCAGAGAACCACTATGACAGGACCTGATGAGTCCAGCTGAGATGGAGGACTCAAA 611
Db 861 AGCAGAGAACCACTATGACAGGACCTGATGAGTCCAGCTGAGATGGAGGACTCAAA 920
Qy 612 GCCACACCCAGTCTCCAGTGTAGCCCTTACTCCAGGCCCCCTTCAAGCCCTGTGAGTACT 671
Db 921 GCCACACCCAGTGTCCAGTGTAGCCCTTACTCCAGGCCCCCTTCAAGCCCTGTGAGTACT 980
Qy 672 CTTTGAAGCTGGGGCATCCGCTGGGCCATCGTGGCCATCGTGTGCTCTCCGTGCTCTG 731
Db 981 CTTTGAAGCTGGGGCATCCGCTGGGCCATCGTGTGCTCTCCGTGCTCTG 1040
Qy 732 CAATGGACTGTGTGTGACCGTGTGCTGGGGGCTTCCGCCCTTGCCTCCCGGTCAA 791
Db 1041 CAATGGACTGTGTGCTGACCGTGTGCTGGGGGCTTCCGCCCTTGCCTCCCGGTCAA 1100
Qy 792 GTTGTGGTAGTGGGATTCAGGCGCCCAACACTTGACTGGCATTTCTGTGGCTTCT 851
Db 1101 GTTGTGGTAGTGGGATTCAGGCGCCCAACACTTGACTGGCATTTCTGTGGCTTCT 1160
Qy 852 AGCTCAGTCAATGCCCTGACCTTTTGGTCACTTCTGAGTACGAGGCCCTGGAGAC 911
Db 1161 AGCTCAGTCAATGCCCTGACCTTTTGGTCACTTCTGAGTACGAGGCCCTGGAGAC 1220
Qy 912 GGGGCTAGGTGCGGGGCCACTTGGCTTCTTGGCAGTACTTGGGTGCGAGGATCGGTGCT 971
Db 1221 GGGGCTAGGTGCGGGGCCACTTGGCTTCTTGGCAGTACTTGGGTGCGAGGATCGGTGCT 1280
Qy 972 GCTGCTACTTGGCGGAGTGCAGTGCAGGCTTCCGTCTCTGTGCTCCGGGCTATGG 1031
Db 1281 GCTGCTACTTGGCGGAGTGCAGTGCAGGCTTCCGTCTCTGTGCTCCGGGCTATGG 1340
Qy 1032 GAAGTCCCTCCCTGGGCGAGCTTCGAGCAGGGGTCTTAGGCTGCTGGCACTGGCAGG 1091
Db 1341 GAAGTCCCTCCCTGGGCGAGCTTCGAGCAGGGGTCTTAGGCTGCTGGCACTGGCAGG 1400
Qy 1092 GCTGGCGCGGCACTGCGGCTTGGCTTCTGAGTGGGAGAAATACGGGGCTCTCCCACTTGTGCT 1151
Db 1401 GCTGGCGCGGCACTGCGGCTTGGCTTCTGAGTGGGAGAAATACGGGGCTCTCCCACTTGTGCT 1460
Qy 1152 GCGCTAGCGGCCACTGAGGGTTCAGCGAGCCCTGGGCTTTCACCTGGGCTCTGGTGAT 1211
Db 1461 GCGCTAGCGGCCACTGAGGGTTCAGCGAGCCCTGGGCTTTCACCTGGGCTCTGGTGAT 1520
Qy 1212 GATGAATCTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1271
Db 1521 GATGAATCTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1580
Qy 1272 GCGCGGGGGGACTTTTGGAGCCGTGTGGGACTGGGCCATGGTGGAGGACGCTGGGCTTGGCT 1331
Db 1581 GCGCGGGGGGACTTTTGGAGCCGTGTGGGACTGGGCCATGGTGGAGGACGCTGGGCTTGGCT 1640
Qy 1332 CATCTTGGCAGACGGGCTCTCTACTGTCCGCTGGGCTTCTCTAGCTTGGCTTGGCTTGGCT 1391
Db 1641 CATCTTGGCAGACGGGCTCTCTACTGTCCGCTGGGCTTCTCTAGCTTGGCTTGGCTTGGCT 1700

QY 1392 GGGGCTCTTCCCTGTCACGCCGAGCGCTCAAGTCTGTCTGCTGGTGGTGGCTGCCCT 1451
DB 1701 GGGGCTCTTCCCTGTACGCCGAGCGCTCAAGTCTGTCTGCTGGTGGTGGCTGCCCT 1760
QY 1452 GCCTGCTGCTCAACCCACTGCTGTACCTGTCTTCAACCCCACTTCCGGGATGACCT 1511
DB 1761 GCCTGCTGCTCAACCCACTGCTGTACCTGTCTTCAACCCCACTTCCGGGATGACCT 1820
QY 1512 TCGGGGCTTGGGCGGCGGCGGAGGAGTCAAGGCGCTAGGCGCTATGCTATGCTGGCGCGGGA 1571
DB 1821 TCGGGGCTTGGGCGGCGGCGGAGGAGTCAAGGCGCTAGGCGCTATGCTATGCTGGCGCGGGA 1880
QY 1572 GCTGAGAGAGCTCTCTGTGATTCTACCCAGGCGCTGTAGCTTCTCTGTATGTGGATCT 1631
DB 1881 GCTGAGAGAGCTCTCTGTGATTCTACCCAGGCGCTGTAGCTTCTCTGTATGTGGATCT 1940
QY 1632 CATCTGGAAGCTTCTGAAGCTGGGCGGCGGCGCTGGGCTGGAGACTATGCTTCCCGCTC 1691
DB 1941 CATCTGGAAGCTTCTGAAGCTGGGCGGCGGCGCTGGGCTGGAGACTATGCTTCCCGCTC 2000
QY 1692 AGTGACCTCATCTCTGTAGAGCCAGGCGGCGGCGGCGCTGGAGGCGAGCCATTGTGT 1751
DB 2001 AGTGACCTCATCTCTGTAGAGCCAGGCGGCGGCGGCGCTGGAGGCGAGCCATTGTGT 2060
QY 1752 AGAGCCAGAGGGGAGCACTTTGGGAACCCCAACCTCCATGGATGGAGAACTGCTGCT 1811
DB 2061 AGAGCCAGAGGGGAGCACTTTGGGAACCCCAACCTCCATGGATGGAGAACTGCTGCT 2120
QY 1812 GAGGCGAGAGGAGTCTAGCCAGCAGGTGGAGGCTTGTTCAGGGGCTGGCGGCTTTTCAGCC 1871
DB 2121 GAGGCGAGAGGAGTCTAGCCAGCAGGTGGAGGCTTGTTCAGGGGCTGGCGGCTTTTCAGCC 2180
QY 1872 CTCTGGCTTGGCTTTGCTTTCACAGTG 1899
DB 2181 CTCTGGCTTGGCTTTGCTTTCACAGTG 2208

RESULT 10

ADN02247
ID ADN02247 standard; cDNA; 2208 BP.
XX
AC ADN02247;
XX
DT 17-JUN-2004 (first entry)
XX
DE Human partial large G protein-coupled receptor, LGR6, CDS #2.
XX
KW Human; ss; large G protein-coupled receptor 6; LGR6; heart failure;
KW ischaemic heart disease; myocardial infarction; hypertension;
KW pericarditis; atherosclerosis; hepatic failure; viral hepatitis; cancer;
KW intracerebral haemorrhage; brain abscess; Alzheimer's disease;
KW Pick disease; Parkinson's disease; Huntington's disease;
KW amyotrophic lateral sclerosis; inflammatory bowel disease;
KW Crohn's disease; benign prostatic hypertrophy;
KW systemic lupus erythematosus.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
XX CDS 1..2208
FT /*tag= a
FT /partial
FT /note= "No start or stop codon shown"
FT /product= "partial LGR6"
XX
XX US2003166047-A1.
XX
XX 04-SEP-2003.
XX
XX 08-MAY-2001; 2001US-00851595.
XX
XX 06-MAY-1999; 99US-0132896P.
XX

PR 08-MAY-2000; 2000US-00566588.
XX (MILL-) MILLENNIUM PHARM INC.
XX

PI Gu W;

XX WPI; 2003-898067/82.
DR P-PSDB; ADN02246.

XX New isolated LGR6 nucleic acids and polypeptides, useful for diagnosing
PT and treating LGR6-mediated disorders, such as myocardial infarction,
PT hypertension, atherosclerosis, viral hepatitis, cancer and/or Alzheimer's
PT disease.

XX Claim 1; SEQ ID NO 9; 145pp; English.

XX The invention relates to an isolated LGR6 (large G protein-coupled
CC receptor) nucleic acid molecule comprising the mouse and human cDNA
CC sequences appearing as ADN02239, ADN02241, ADN02242, ADN02244, ADN02245,
CC ADN02247, ADN02248 and ADN02250. Also included are a host cell or a non-
CC human mammalian host cell containing an LGR6 nucleic acid (and producing
CC LGR6), an isolated LGR6 polypeptide appearing as ADN02240, ADN02243,
CC ADN02246 and ADN02249, an antibody which selectively binds to LGR6, a
CC compound which selectively binds to LGR6, or hybridises to LGR6 nucleic
CC acid, and instructions for use) a method for detecting LGR6 nucleic acid
CC in a sample, a method for identifying a compound which binds to LGR6, a
CC method for modulating the activity of LGR6 and a method for identifying a
CC compound that modulates the activity of LGR6. The methods and
CC compositions of the present invention are useful for diagnosing and
CC treating large G-protein coupled receptor 6 (LGR6)-mediated or -related
CC disorders, such as heart failure, ischaemic heart disease, myocardial
CC infarction, hypertension, pericarditis, intracerebral haemorrhage, brain abscess,
CC viral hepatitis, cancer, intracerebral haemorrhage, Alzheimer's disease,
CC Alzheimer's disease, Pick disease, Parkinson's disease, Huntington's
CC disease, amyotrophic lateral sclerosis, inflammatory bowel disease,
CC Crohn's disease, benign prostatic hypertrophy and systemic lupus
CC erythematosus. The present sequence is a partial human CDS for LGR6.

XX Sequence 2208 BP; 396 A; 710 C; 631 G; 471 T; 0 U; 0 Other;

Query Match 92.4%; Score 1754.4; DB 11; Length 2208;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1766; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

QY 132 AGCGGCTTCCAAACACACACCGCATCTGGGAAATTGGAGCTNGACACCTTCAGCCAGCTGAG 191
DB 442 ATCGGCTTCCAAACACACACCGCATCTGGGAAATTGGAGCT-GACACCTTCAGCCAGCTGAG 500
QY 192 CTCCTGCGAAGCCCTGGATCTTAGCTGGAAAGCCATCCGGTCCATCCACCTGAGGCGCTT 251
DB 501 CTCCTGCGAAGCCCTGGATCTTAGCTGGAAAGCCATCCGGTCCATCCACCTGAGGCGCTT 560
QY 252 CTCACACCTTCACCTCCCTGGTCAAGCTGGACCTGACAGACAAACAGCTGACCACTGCC 311
DB 561 CTCACACCTTCACCTCCCTGGTCAAGCTGGACCTGACAGACAAACAGCTGACCACTGCC 620
QY 312 CTTGGCTGGAATTGGGGGCTTGATGATCTGAAGCTCAAAAGGAACTTGTCTCTCTCCCA 371
DB 621 CTTGGCTGGAATTGGGGGCTTGATGATCTGAAGCTCAAAAGGAACTTGTCTCTCTCCCA 680
QY 372 GGCCTTCTCCAAGGACAGTTTCCCAAACTGAGGATCCTCGAGGTGCTTTATGCTTACCA 431
DB 681 GGCCTTCTCCAAGGACAGTTTCCCAAACTGAGGATCCTCGAGGTGCTTTATGCTTACCA 740
QY 432 GTGCTGTCCCTATGGGATGTGTCAGCTTCTTCAAGGCTCTGGGCGAGTGGAGGCTGA 491
DB 741 GTGCTGTCCCTATGGGATGTGTCAGCTTCTTCAAGGCTCTGGGCGAGTGGAGGCTGA 800
QY 492 AGACCTTCACCTTGATGATGAGGAGTCTTCAAAAGGCGCTTGGGCGCTCTCTTGCACACA 551
DB 801 AGACCTTCACCTTGATGATGAGGAGTCTTCAAAAGGCGCTTGGGCGCTCTCTTGCACACA 860

```
QY 552 AGCAGAGAACCACTATGACAGGACTCTGGATGAGCTCCAGCTGGAGATGGAGACTCAAA 611
Db 861 AGCAGAGAACCACTATGACAGGACTCTGGATGAGCTCCAGCTGGAGATGGAGACTCAAA 920
QY 612 GCCACACCCAGTGTCCAGTGTAGCCCTACTTCCAGGCCCTTCAAGCCCTGTGAGTACT 671
Db 921 GCCACACCCAGTGTCCAGTGTAGCCCTACTTCCAGGCCCTTCAAGCCCTGTGAGTACT 980
QY 672 CTTTGAAGCTGGGGCATCCGCTTGGCCGTGTGGCCCATCGTGTGCTCTCCGTCCTCTG 731
Db 981 CTTTGAAGCTGGGGCATCCGCTTGGCCGTGTGGCCCATCGTGTGCTCTCCGTCCTCTG 1040
QY 732 CAATGGACTGTGCTGTACCGTGTTCGTGGCGGGCTGCCCCCTGCCCCCGGTCAA 791
Db 1041 CAATGGACTGTGCTGTACCGTGTTCGTGGCGGGCTGCCCCCTGCCCCCGGTCAA 1100
QY 792 GTTTCGTGTAGTGTGATTTGACGGCGCCAAACCTTTGACTGTGCACTTCTGTGGCTTCT 851
Db 1101 GTTTCGTGTAGTGTGATTTGACGGCGCCAAACCTTTGACTGTGCACTTCTGTGGCTTCT 1160
QY 852 AGCCTCAGTGCATGCCCTGACCTTTTGGTCACTTCTCTGAGTACGAGAGCCCGCTGGAGAC 911
Db 1161 AGCCTCAGTGCATGCCCTGACCTTTTGGTCACTTCTCTGAGTACGAGAGCCCGCTGGAGAC 1220
QY 912 GGGGCTAGGCTGCCCGGCCACATGGCTTCTTGCCAGTACTTGGGTGCGAGGCATCGGTGCT 971
Db 1221 GGGGCTAGGCTGCCCGGCCACATGGCTTCTTGCCAGTACTTGGGTGCGAGGCATCGGTGCT 1280
QY 972 GCTGCTCACTCTGGCGGAGTGCAGTGCAGCGTCTCCGTCCTCTGTCGCCGGCCATGCG 1031
Db 1281 GCTGCTCACTCTGGCGGAGTGCAGTGCAGCGTCTCCGTCCTCTGTCGCCGGCCATGCG 1340
QY 1032 GAAATGCCCTCCCTGGGCGAGCTTTCGAGCAGGGGTCTTAGGCTGCCCTGGCACTGGCAGG 1091
Db 1341 GAAATGCCCTCCCTGGGCGAGCTTTCGAGCAGGGGTCTTAGGCTGCCCTGGCACTGGCAGG 1400
QY 1092 GCTGCGCGCGCACTTGCCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1151
Db 1401 GCTGCGCGCGCACTTGCCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1460
QY 1152 GCCCTACCGGCACCTGAGGCTCAGCCAGCAGCCCTGGGCTTACCGTGGCCCTGGTGTAT 1211
Db 1461 GCCCTACCGGCACCTGAGGCTCAGCCAGCAGCCCTGGGCTTACCGTGGCCCTGGTGTAT 1520
QY 1212 GATGAACCTCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1271
Db 1521 GATGAACCTCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1580
QY 1272 GCCCGGGGCGACTTTGAGGCGCTGTGGGACTGCGCCCATGTGTGAGGCACTGGCCCTGGCT 1331
Db 1581 GCCCGGGGCGACTTTGAGGCGCTGTGGGACTGCGCCCATGTGTGAGGCACTGGCCCTGGCT 1640
QY 1332 CATCTTCGAGAGCGGGCTCTCTACTGTCCGCTGGGCTTCTCAGCTTGGCCCTCATGCT 1391
Db 1641 CATCTTCGAGAGCGGGCTCTCTACTGTCCGCTGGGCTTCTCAGCTTGGCCCTCATGCT 1700
QY 1392 GGGGCTCTTCTCTGCTCAGCCCGAGCCGTCAAGTCTGTCTGCTGCTGCTGCTGCTGCTGCT 1451
Db 1701 GGGGCTCTTCTCTGCTCAGCCCGAGCCGTCAAGTCTGTCTGCTGCTGCTGCTGCTGCTGCT 1760
QY 1452 GCCTGCTGCTCAACCCACTGCTGTACCTGTCTTTCAACCCCACTTCCGGGATGACCT 1511
Db 1761 GCCTGCTGCTCAACCCACTGCTGTACCTGTCTTTCAACCCCACTTCCGGGATGACCT 1820
QY 1512 TCGGCGGCTTGGGCGCGGAGGGGACTCAGGGCCCTTAGGCTATGTGTGGCGCGGGGA 1571
Db 1821 TCGGCGGCTTGGGCGCGGAGGGGACTCAGGGCCCTTAGGCTATGTGTGGCGCGGGGA 1880
QY 1572 GCTGAGAGAGCTCTCTGTGATTTACCCAGGCCCTGGTGTAGCTTCTCTGATGTGATCT 1631
Db 1881 GCTGAGAGAGCTCTCTGTGATTTACCCAGGCCCTGGTGTAGCTTCTCTGATGTGATCT 1940
QY 1632 CATTCTGAAGCTTCTGAAGCTGGGCGGCGGCCCCCTGGGCTGGAGACTTATGGCTTCCCCCTC 1691
```

```
Db 1941 CATTCTGAAGCTTCTGAAGCTGGGCGGCCCCCTGGGCTGGAGACTTATGCTTCCCCCTC 2000
QY 1692 AGTGACCCCTCATCTCCTGTGTCAGCAGCCAGGGGCCCCCAGGCTGGAGGCGAGCATTTGTGT 1751
Db 2001 AGTGACCCCTCATCTCCTGTGTCAGCAGCCAGGGGCCCCCAGGCTGGAGGCGAGCATTTGTGT 2060
QY 1752 AGAGCCAGAGGGGAACCACTTTGGGAACCCCAACCCCTCCATGTGATGAGAACTGCTGCT 1811
Db 2061 AGAGCCAGAGGGGAACCACTTTGGGAACCCCAACCCCTCCATGTGATGAGAACTGCTGCT 2120
QY 1812 GAGGCGCAGAGGATCTACGCCAGCAGGTGGAGGCTTGTTCAGGGGGTGGCGGCTTTCAGCC 1871
Db 2121 GAGGCGCAGAGGATCTACGCCAGCAGGTGGAGGCTTGTTCAGGGGGTGGCGGCTTTCAGCC 2180
QY 1872 CTCTGGCTTGGCCCTTTGCTTTCACACGCTG 1899
Db 2181 CTCTGGCTTGGCCCTTTGCTTTCACACGCTG 2208

RESULT 11
ADK19412
ID ADK19412 standard; cDNA; 2208 BP.
XX
AC ADK19412;
XX
DT 17-JUN-2004 (first entry)
XX
DE ORF of partial cDNA encoding human LGR6 polypeptide #2.
XX
KW G-protein coupled receptor; GPCR; large G-protein coupled receptor 6;
LGR6; weight disorder; neural disorder; endocrine disorder;
KW cardiovascular disorder; anorectic; anabolic; immunomodulator;
KW neuroprotective; nootropic; antiparkinsonian; anti-HIV; anticonvulsant;
KW antihypertensive; osteopathic; cardiac; vasotropic; hypotensive;
KW antiarteriosclerotic; antiarrhythmic; human; gene; ss.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
CDS 1..2208
FT /*tag= a
FT /partial
FT /product= "LGR6 polypeptide #2"
FT /note= "This sequence lacks both start and stop codons"
XX
PN US2004058377-A1.
XX
PD 25-MAR-2004.
XX
PF 18-SEP-2003; 2003US-00664667.
XX
PR 06-MAY-1999; 99US-0132896P.
PR 08-MAY-2000; 2000US-00566588.
PR 08-MAY-2001; 2001US-00851595.
XX
PA (MILL-) MILLENNIUM PHARM INC.
XX
XX Gu W;
XX
XX WPI; 2004-268789/25.
XX P-PDB; ADK19411.
XX
PT New large G-protein coupled receptor 6 nucleic acid molecules and
polypeptides, useful for diagnosing, preventing or treating diseases
PT associated with aberrant nucleic acid or protein activity, e.g. obesity,
PT anorexia or cachexia.
XX
PS Claim 2; SEQ ID NO 9; 145pp; English.
XX
CC The present invention relates to the isolation of novel members of the G-
protein coupled receptor (GPCR) family designated as large G-protein
coupled receptor 6 (LGR6), and the polynucleotide sequences encoding
```


CC them. The invention also discloses LGR6 fusion proteins, antigenic
CC peptides, anti-LGR6 antibodies, recombinant expression vectors, host
CC cells, and non-human transgenic animals in which an LGR6 gene has been
CC introduced or disrupted. The LGR6 polynucleotide sequences, polypeptide
CC sequences, and antibodies are useful for diagnosing, treating or
CC preventing diseases associated with aberrant LGR6 expression or activity,
CC such as weight disorders (e.g. obesity, anorexia nervosa or cachexia),
CC neural disorders (e.g. central nervous system disorder, including
CC Alzheimer's disease, Parkinson's disease, AIDS-related dementia, multiple
CC sclerosis and Huntington's disease), endocrine disorders (e.g.
CC hypothyroidism, hyperthyroidism or acromegaly), or cardiovascular
CC disorders (e.g. ischemia-reperfusion injury, coronary artery disease,
CC myocardial infarction, arrhythmia, atherosclerosis, hypertension or
CC congestive heart failure). The LGR6 polynucleotide and polypeptide
CC sequences are also useful as targets for developing modulating agents
CC that regulate a variety of cellular processes, e.g. neural and endocrine
CC processes and thermogenesis. They can be used in screening assays (e.g.
CC chromosome mapping, tissue typing or in forensic biology), or in
CC predictive medicine (e.g. diagnostic assays, prognostic assays,
CC monitoring clinical trials or pharmacogenomics). The present sequence
CC encodes a human LGR6 polypeptide.
XX
SQ Sequence 2208 BP; 396 A; 710 C; 631 G; 471 T; 0 U; 0 Other;

Query Match 92.4%; Score 1754.4; DB 12; Length 2208;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1766; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

QY 132 AGCGCCCTCCACACACACCGCATCTGGGAAATTGGAGCTNGACACCTTTCAGCCAGCTGAG 191
DB 442 ATGCGCCCTCCACACACACCGCATCTGGGAAATTGGAGCT-GACACCTTTCAGCCAGCTGAG 500

QY 192 CTCCTGCAAGCCCTGGATCTTAGCTGGAAGCCCATCCGGTCCATCCACCTGAGGCCCTT 251
DB 501 CTCCTGCAAGCCCTGGATCTTAGCTGGAAGCCCATCCGGTCCATCCACCTGAGGCCCTT 560

QY 252 CTCACCCCTGCACCTCCCTGGTCAAGCTGACCTGACACACACACAGCTGACACACATGCC 311
DB 561 CTCACCCCTGCACCTCCCTGGTCAAGCTGACCTGACACACACAGCTGACACACATGCC 620

QY 312 CTTGCTGGACTTGGGGCTTGGATCATCTGAAGCTCAAGGGAACCTTGTCTCTCCCA 371
DB 621 CTTGCTGGACTTGGGGCTTGGATCATCTGAAGCTCAAGGGAACCTTGTCTCTCCCA 680

QY 372 GGCCTTCTCCAGGACAGTTTCCCAAACTGAGATCTCGAGGTGCTTATGCTTACCA 431
DB 681 GGCCTTCTCCAGGACAGTTTCCCAAACTGAGATCTCGAGGTGCTTATGCTTACCA 740

QY 432 GTGCTGCTCCCTATGGGATGTGTCAGCTTCTTCAAGGCCCTCTGGGCAAGTGGAGGCTGA 491
DB 741 GTGCTGCTCCCTATGGGATGTGTCAGCTTCTTCAAGGCCCTCTGGGCAAGTGGAGGCTGA 800

QY 492 AGACCTTCCACCTTGATGATGAGGAGTCTTCAAAAGGCCCTTGGGCTCTCTTGGCAGACA 551
DB 801 AGACCTTCCACCTTGATGATGAGGAGTCTTCAAAAGGCCCTTGGGCTCTCTTGGCAGACA 860

QY 552 AGCAGAGAACCATATGACACAGGACCTTGGATGAGCTCCAGCTGGAGATGAGGAGCTCAAA 611
DB 861 AGCAGAGAACCATATGACACAGGACCTTGGATGAGCTCCAGCTGGAGATGAGGAGCTCAAA 920

QY 612 GCCACACCCAGTGTCCAGTGTAGCCCTTCTCCAGGCCCTTCAAGCCCTGTGAGTACCT 671
DB 921 GCCACACCCAGTGTCCAGTGTAGCCCTTCTCCAGGCCCTTCAAGCCCTGTGAGTACCT 980

QY 672 CTTTGAAGAGCTGGGCAATCCGCTGGCGGTGGGCCCATCGTGTGCTCTCCGCTGCTG 731
DB 981 CTTTGAAGAGCTGGGCAATCCGCTGGCGGTGGGCCCATCGTGTGCTCTCCGCTGCTG 1040

QY 732 CAATGGAGTGTGCTGTGACCGTGTTCGCTGGGGGGCTGCCCTCCCTGCCCCCGGTCAA 791
DB 1041 CAATGGAGTGTGCTGTGACCGTGTTCGCTGGGGGGCTGCCCTCCCTGCCCCCGGTCAA 1100

QY 792 GTTTGTGGTAGTGGATTTGACGGCGGCAACACCTTGACTGGCATTTCTGTGCGCTTCT 851

DB 1101 GTTTGTGGTAGTGGATTTGACGGCGCAACACCTTGACTGCGATTTCTGTGCGCTTCT 1160

QY 852 AGCCTCAGTCAGATCCCTGACCTTTGGTCACTTCTCTGAGTACGAGAGCCCTGGGAGAC 911

DB 1161 AGCCTCAGTCAGATCCCTGACCTTTTGGTCACTTCTCTGAGTACGAGAGCCCTGGGAGAC 1220

QY 912 GGGCTAGGCTGCGCGGCCCACTGGCTTCTCTGGCAGTACTTGGGTGCGAGGACATCGGTGCT 971

DB 1221 GGGCTAGGCTGCGCGGCCCACTGGCTTCTCTGGCAGTACTTGGGTGCGAGGACATCGGTGCT 1280

QY 972 GCTGCTCACTCTGCGCCGACAGTGCAGCGTCTCCGCTCTCTGTGTCTCCGGGCTATGG 1031

DB 1281 GCTGCTCACTCTGCGCCGACAGTGCAGCGTCTCCGCTCTCTGTGTCTCCGGGCTATGG 1340

QY 1032 GAATGCCCTCTCTGCGGACAGCTTCAGAGAGGGGTCTTAGGCTGCGCTGGACATGGCAGG 1091

DB 1341 GAATGCCCTCTCTGCGGACAGCTTCAGAGAGGGGTCTTAGGCTGCGCTGGACATGGCAGG 1400

QY 1092 GCTGGCGCGCAGCTGCGCCCTGGCTCAGTGGGAGATACGGGGCTTCCGACATCTGCCT 1151

DB 1401 GCTGGCGCGCAGCTGCGCCCTGGCTCAGTGGGAGATACGGGGCTTCCGACATCTGCCT 1460

QY 1152 GCCCTACGCGCCACTGAGGGTCAAGCCAGCAGCCCTGGGCTTCAACGCTGGGCCCTTGGTAT 1211

DB 1461 GCCCTACGCGCCACTGAGGGTCAAGCCAGCAGCCCTGGGCTTCAACGCTGGGCCCTTGGTAT 1520

QY 1212 GATGAATCTCTCTGTTTCTGTCGTGGCGGTCTCATCAAACTGTAATCAAACTGTAATCAAACT 1271

DB 1521 GATGAATCTCTCTGTTTCTGTCGTGGCGGTCTCATCAAACTGTAATCAAACTGTAATCAAACT 1580

QY 1272 GCCGCGGGGCGACTTTGAGGCGGTGGGACTGCGGCATGGTGGAGGACGCTGGGCTGGCT 1331

DB 1581 GCCGCGGGGCGACTTTGAGGCGGTGGGACTGCGGCATGGTGGAGGACGCTGGGCTGGCT 1640

QY 1332 CATCTTCGACAGAGGGGCTCTTACTGTCGCGTGGGCTTCTCTCAAGCTTGGGCTCCATGCT 1391

DB 1641 CATCTTCGACAGAGGGGCTCTTACTGTCGCGTGGGCTTCTCTCAAGCTTGGGCTCCATGCT 1700

QY 1392 GGGCTCTTCTCCTGTACGCCCGAGGCGCTCAAGTCTGTCTGTGTGTGTGTGTGTGTGTGTGT 1451

DB 1701 GGGCTCTTCTCCTGTACGCCCGAGGCGCTCAAGTCTGTCTGTGTGTGTGTGTGTGTGTGTGT 1760

QY 1452 GCTGCTGCTCTCAACCCACTGCTGTACCTTCTTCAACCCCACTTCCGCGGATGACCT 1511

DB 1761 GCTGCTGCTCTCAACCCACTGCTGTACCTTCTTCAACCCCACTTCCGCGGATGACCT 1820

QY 1512 TCGGCGGCTTTCGGCGCCCGCAGGAGACTCAGGGCCCTTAGCCCTATGCTGCGGCGCGGGA 1571

DB 1821 TCGGCGGCTTTCGGCGCCCGCAGGAGACTCAGGGCCCTTAGCCCTATGCTGCGGCGCGGGA 1880

QY 1572 GCTGGAGAGAGCTCTCTGTGATTTACCCAGGCGCTTGGTGTGCTGTGTGTGTGTGTGTGTGT 1631

DB 1881 GCTGGAGAGAGCTCTCTGTGATTTACCCAGGCGCTTGGTGTGCTGTGTGTGTGTGTGTGTGT 1940

QY 1632 CATCTTGAAGCTTCTGAAGCTGGGCGGCCCTTGGGCTGGAGACCTATGGCTTCCCTC 1691

DB 1941 CATCTTGAAGCTTCTGAAGCTGGGCGGCCCTTGGGCTGGAGACCTATGGCTTCCCTC 2000

QY 1692 AGTGACCTCATCTCTCTGTGAGAGCCAGGCGGCCCTCAGGCTGGAGGCGCAGCCATTTGCT 1751

DB 2001 AGTGACCTCATCTCTCTGTGAGAGCCAGGCGGCCCTCAGGCTGGAGGCGCAGCCATTTGCT 2060

QY 1752 AGAGCCAGAGGGGAAACCACTTTGGGAAACCCCAACCTTCCATGATGAGAGAACTGTGCT 1811

DB 2061 AGAGCCAGAGGGGAAACCACTTTGGGAAACCCCAACCTTCCATGATGAGAGAACTGTGCT 2120

QY 1812 GAGGCGAGAGGGATCTACGCCAGCTGTGAGGCTGTGAGGGGTGCGGCTTTCAGCC 1871

DB 2121 GAGGCGAGAGGGATCTACGCCAGCTGTGAGGCTGTGAGGGGTGCGGCTTTCAGCC 2180

QY 1872 CTCTGGCTTGGCTTTCAGCATCTG 1899

Db 1461 GCCCTACGCCACCTGAGGGGTACGCCAGCAGCCCTTGGGCTTCAACCGTGGCCCTGGTGAT 1520
 Qy 1212 GATGAATCCTTCTGTTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1271
 Db 1521 GATGAATCCTTCTGTTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1580
 Qy 1272 GCCCGGGGCGACTTTGAGGGCGTGTGGGACTGCGCCATGCTGAGGACACGTGGCCCTGGCT 1331
 Db 1581 GCCCGGGGCGACTTTGAGGGCGTGTGGGACTGCGCCATGCTGAGGACACGTGGCCCTGGCT 1640
 Qy 1332 CATCTTGGAGAGCGGCTCTCTACTGCTCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1391
 Db 1641 CATCTTGGAGAGCGGCTCTCTACTGCTCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1700
 Qy 1392 GGGCGCTTCCCTGTCAGCCCGGCGGCGTCAAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1451
 Db 1701 GGGCGCTTCCCTGTCAGCCCGGCGGCGTCAAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1760
 Qy 1452 GCTGCTGCTGCTCAACCCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1511
 Db 1761 GCTGCTGCTGCTCAACCCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1820
 Qy 1512 TCGCGGCTTGGCCCGCGCAGGGGACTCAGGCGCCCTAGCCTATGCTGCGGCGCGGGA 1571
 Db 1821 TCGCGGCTTGGCCCGCGCAGGGGACTCAGGCGCCCTAGCCTATGCTGCGGCGCGGGA 1880
 Qy 1572 GCTGAGAGAGCTCCTGCTGATTTACCCAGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1631
 Db 1881 GCTGAGAGAGCTCCTGCTGATTTACCCAGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1940
 Qy 1632 CATCTGGAAGCTTCTGAAGCTGCGCGCGCCCTGCGGCTGAGAGCTATGCTTCCCTC 1691
 Db 1941 CATCTGGAAGCTTCTGAAGCTGCGCGCGCCCTGCGGCTGAGAGCTATGCTTCCCTC 2000
 Qy 1692 AGTACCTCTATCTCTCTGACAGCCAGGCGCGCCCTGAGGCTGAGGAGCAGCCATTTGCT 1751
 Db 2001 AGTACCTCTATCTCTCTGACAGCCAGGCGCGCCCTGAGGCTGAGGAGCAGCCATTTGCT 2060
 Qy 1752 AGAGCCAGAGGGGACCACTTTGGGAGCCGCCACCCCTCCATGATGAGAGCTGCTGCT 1811
 Db 2061 AGAGCCAGAGGGGACCACTTTGGGAGCCGCCACCCCTCCATGATGAGAGCTGCTGCT 2120
 Qy 1812 GAGGCGAGAGGGATCTAGCCAGCAGGTGGAGGCTTGTGAGGGGTGGCGGCTTTCAGCC 1871
 Db 2121 GAGGCGAGAGGGATCTAGCCAGCAGGTGGAGGCTTGTGAGGGGTGGCGGCTTTCAGCC 2180
 Qy 1872 CTCTGGCTTGGCTTTGCTTTCACAGTG 1899
 Db 2181 CTCTGGCTTGGCTTTGCTTTCACAGTG 2208

RESULT 13

ID ADN02245 standard; cDNA; 2711 BP.
 AC ADN02245;
 XX ADN02245;
 DT 17-JUN-2004 (first entry)
 XX Human partial large G protein-coupled receptor, LGR6, cDNA #2.
 DE Human; ss; large G protein-coupled receptor 6; LGR6; heart failure;
 KW ischaemic heart disease; myocardial infarction; hypertension;
 KW pericarditis; atherosclerosis; hepatic failure; viral hepatitis; cancer;
 KW intracerebral haemorrhage; brain abscess; Alzheimer's disease;
 KW Pick disease; Parkinson's disease; Huntington's disease;
 KW amyotrophic lateral sclerosis; inflammatory bowel disease;
 KW Crohn's disease; benign prostatic hypertrophy;
 KW systemic lupus erythematosus.
 OS Homo sapiens.
 XX
 XX
 FH Key Location/Qualifiers

CDS 1. .2211
 FT /*tag= a
 FT /partial
 FT /note= "No start codon shown"
 FT /product= "Partial LGR6"
 XX
 PN US2003166047-A1.
 XX
 PD 04-SEP-2003.
 XX
 PF 08-MAY-2001; 2001US-00851595.
 XX
 PR 06-MAY-1999; 99US-0132896P.
 PR 08-MAY-2000; 2000US-00566588.
 XX
 PA (MILL-) MILLENNIUM PHARM INC.
 XX
 FI Gu W;
 XX
 WI WPI: 2003-898067/82.
 P-PSDB; ADN02246.
 XX
 PT New isolated LGR6 nucleic acids and polypeptides, useful for diagnosing and treating LGR6-mediated disorders, such as myocardial infarction, hypertension, atherosclerosis, viral hepatitis, cancer and/or Alzheimer's disease.
 XX
 PT Claim 1; SEQ ID NO 7; 145pp; English.
 XX
 CC The invention relates to an isolated LGR6 (large G protein-coupled receptor) nucleic acid molecule comprising the mouse and human cDNA sequences appearing as ADN02239, ADN02241, ADN02242, ADN02244, ADN02245, ADN02247, ADN02248 and ADN02250. Also included are a host cell or a non-human mammalian host cell containing an LGR6 nucleic acid (and producing LGR6), an isolated LGR6 polypeptide appearing as ADN02240, ADN02243, ADN02246 and ADN02249, an antibody which selectively binds to LGR6, a method for detecting the presence of LGR6 in a sample, a kit (comprising a compound which selectively binds to LGR6, or hybridises to LGR6 nucleic acid, and instructions for use), a method for detecting LGR6 nucleic acid in a sample, a method for identifying a compound which binds to LGR6, a method for modulating the activity of LGR6 and a method for identifying a compound that modulates the activity of LGR6. The methods and compositions of the present invention are useful for diagnosing and treating large G-protein coupled receptor 6 (LGR6)-mediated or -related disorders, such as heart failure, ischaemic heart disease, myocardial infarction, hypertension, pericarditis, atherosclerosis, hepatic failure, viral hepatitis, cancer, intracerebral haemorrhage, brain abscess, Alzheimer's disease, Pick disease, Parkinson's disease, Huntington's disease, amyotrophic lateral sclerosis, inflammatory bowel disease, Crohn's disease, benign prostatic hypertrophy and systemic lupus erythematosus. The present sequence is a partial human cDNA for LGR6.
 XX
 SQ Sequence 2711 BP; 530 A; 837 C; 740 G; 604 T; 0 U; 0 Other;
 Query Match 92.4%; Score 1754.4; DB 11; Length 2711;
 Best Local Similarity 99.9%; Pred. No. 0;
 Matches 1766; Conservative 0; Mismatches 1; Indels 1; Gaps 1;
 Qy 132 AGCGGCTCCAAACACACCCGCTATGGGAAATTGGAGCTNGACACCTTACGCCAGCTGAG 191
 Db 442 ATCGGCTCCAAACACACCCGCTATGGGAAATTGGAGCT-GACACCTTACGCCAGCTGAG 500
 Qy 192 CTCCCTCAGACCTTGGATCTTAGCTGGAAGCCATCGGCTCCATCCACCTGAGGCTT 251
 Db 501 CTCCCTCAGACCTTGGATCTTAGCTGGAAGCCATCGGCTCCATCCACCTGAGGCTT 560
 Qy 252 CTCCACCTCCTGCTCCCTGCTCAAGCTGAGCTGACAGACACCCAGCTGACACCTGCC 311
 Db 561 CTCCACCTCCTGCTCCCTGCTCAAGCTGAGCTGACAGACACCCAGCTGACACCTGCC 620
 Qy 312 CTGGGCTGGAAGCTTGGGGGCTTGTATGATCTGAAGCTCAAGGGAACTTGTCTCTCTCCA 371
 Db 621 CTGGGCTGGAAGCTTGGGGGCTTGTATGATCTGAAGCTCAAGGGAACTTGTCTCTCTCCA 680

```
QY 372 GGCCTTCTCAAAGGACAGTTTCCCAAAACTGAGGATCCTGGAGGTGCCTTATGCCTACCA 431
Db 681 GGCCTTCTCAAAGGACAGTTTCCCAAAACTGAGGATCCTGGAGGTGCCTTATGCCTACCA 740
QY 432 GTGCTGTCCCTATGGATGTGGCAGCTTCTTCAAGGCTCTGGCAGTGGAGGCTGA 491
Db 741 GTGCTGTCCCTATGGATGTGGCAGCTTCTTCAAGGCTCTGGCAGTGGAGGCTGA 800
QY 492 AGACCTTTCACCTTGATGAGGAGTCTTCAAAAAGGCCCTTGGGCTCTCTTGGCCAGACA 551
Db 801 AGACCTTTCACCTTGATGAGGAGTCTTCAAAAAGGCCCTTGGGCTCTCTTGGCCAGACA 860
QY 552 AGCAGAGAACCACTATGACACGGA CTTGGATGAGTCCAGCTCGAGATGGAGGATCAAA 611
Db 861 AGCAGAGAACCACTATGACACGGA CTTGGATGAGTCCAGCTCGAGATGGAGGATCAAA 920
QY 612 GCCACACCCCACTGTCAGTGTAGCCCTACTCCAGGCCCTTCAAGCCCTGTGAGTACT 671
Db 921 GCCACACCCCACTGTCAGTGTAGCCCTACTCCAGGCCCTTCAAGCCCTGTGAGTACT 980
QY 672 CTTTGAAGACTGGGCGATCCGCTGGCCGTGGCCGCAATCGTGTGTCTCCGTGCTGTG 731
Db 981 CTTTGAAGACTGGGCGATCCGCTGGCCGTGGCCGCAATCGTGTGTCTCCGTGCTGTG 1040
QY 732 CAATGGACTGTGTGCTGACCGTGTTCGCTGGCGGCGCTGCCCTTGCCTCCCGGTCAA 791
Db 1041 CAATGGACTGTGTGCTGACCGTGTTCGCTGGCGGCGCTGCCCTTGCCTCCCGGTCAA 1100
QY 792 GTTGTGTGATGTGATGTGAGGGCCCAACCTTGACTGAGCATTTCTGTGGCTTCT 851
Db 1101 GTTGTGTGATGTGATGTGAGGGCCCAACCTTGACTGAGCATTTCTGTGGCTTCT 1160
QY 852 AGCCTCAGTCGATGCCCTTGGCTTGGTCTGAGTACGAGCCCGCTGGAGAC 911
Db 1161 AGCCTCAGTCGATGCCCTTGGCTTGGTCTGAGTACGAGCCCGCTGGAGAC 1220
QY 912 GGGGCTAGGCTGCCGGGCACTGGCTTCTGGCAGTACTTGGGTGGAGGCAATCGGTGT 971
Db 1221 GGGGCTAGGCTGCCGGGCACTGGCTTCTGGCAGTACTTGGGTGGAGGCAATCGGTGT 1280
QY 972 GCTGTCTACTTGGCGGCGAGTGCAGTGCAGCTCTCGTCTCTGTGTCCGGGCTATGG 1031
Db 1281 GCTGTCTACTTGGCGGCGAGTGCAGTGCAGCTCTCGTCTCTGTGTCCGGGCTATGG 1340
QY 1032 GAAATCCCCCTCCCTGGGCGAGGTTTCGAGCAGGGTCTTAGCTGCCCTGGCACTGGCAGG 1091
Db 1341 GAAATCCCCCTCCCTGGGCGAGGTTTCGAGCAGGGTCTTAGCTGCCCTGGCACTGGCAGG 1400
QY 1092 GCTGGCCGCGCACTGCCCTTGGCTCAGTGGGAGAATACGGGGCTTCCCACTCTGCT 1151
Db 1401 GCTGGCCGCGCACTGCCCTTGGCTCAGTGGGAGATACGGGGCTTCCCACTCTGCT 1460
QY 1152 GCCCTACGGGCACTGAGGGTTCAGCAGAGCTTGGGCTTTCACGGTGGCCCTGGTAT 1211
Db 1461 GCCCTACGGGCACTGAGGGTTCAGCAGAGCTTGGGCTTTCACGGTGGCCCTGGTAT 1520
QY 1212 GATGAACCTCTCTCTTCTGCTGTGGTGGCCGGTGCCTACATCAAACTGTACTGTGACT 1271
Db 1521 GATGAACCTCTCTCTTCTGCTGTGGTGGCCGGTGCCTACATCAAACTGTACTGTGACT 1580
QY 1272 GCCCGGGGCGACTTTGAGGGCCGTGTGGGACTGCGCCATGTGTGAGGACAGTGGCTGGCT 1331
Db 1581 GCCCGGGGCGACTTTGAGGGCCGTGTGGGACTGCGCCATGTGTGAGGACAGTGGCTGGCT 1640
QY 1332 CATCTTCGAGACGGGCTCTCTACTGTGTCGGTGGCCCTTCTCAGCTTCGCTCCATGCT 1391
Db 1641 CATCTTCGAGACGGGCTCTCTACTGTGTCGGTGGCCCTTCTCAGCTTCGCTCCATGCT 1700
QY 1392 GGGCTCTTCCCTGTACGCCCGGCGGTCAAGTCTGTCTGTGCTGGTGGTGGCTGCCCT 1451
Db 1701 GGGCTCTTCCCTGTACGCCCGGCGGTCAAGTCTGTCTGTGCTGGTGGTGGCTGCCCT 1760
```

```
QY 1452 GCCTGCTGCTCAACCCACCTGCTGTACTGCTCTTCAACCCCACTTCCGGGATGACT 1511
Db 1761 GCCTGCTGCTCAACCCACCTGCTGTACTGCTCTTCAACCCCACTTCCGGGATGACT 1820
QY 1512 TCGGGGCTTCGGCCCGCGCAGGGGACTCAGGGCCCTAGCCTATGCTGGGCCGGGA 1571
Db 1821 TCGGGGCTTCGGCCCGCGCAGGGGACTCAGGGCCCTAGCCTATGCTGGGCCGGGA 1880
QY 1572 GCTGGAGAAGAGCTCTCTGTGATTCTACCCAGGCCCCCTGGTAGCCTTCTCTGATGTGATCT 1631
Db 1881 GCTGGAGAAGAGCTCTCTGTGATTCTACCCAGGCCCCCTGGTAGCCTTCTCTGATGTGATCT 1940
QY 1632 CATCTGGAAGCTTCTGAAGCTGGGCGGCCCTTGGGCTGGAGACCTATGCTTCCCTC 1691
Db 1941 CATCTGGAAGCTTCTGAAGCTGGGCGGCCCTTGGGCTGGAGACCTATGCTTCCCTC 2000
QY 1692 AGTGACCTCATCTCTGTCAGCAGCCAGGGGCCCAAGGCTGGAGGCGACCATTTGT 1751
Db 2001 AGTGACCTCATCTCTGTCAGCAGCCAGGGGCCCAAGGCTGGAGGCGACCATTTGT 2060
QY 1752 AGAGCCAGAGGGGAACCACTTTGGGAACCCCAACCCCTCCATGGATGGAGAACTCTGCT 1811
Db 2061 AGAGCCAGAGGGGAACCACTTTGGGAACCCCAACCCCTCCATGGATGGAGAACTCTGCT 2120
QY 1812 GAGGCGAGAGGATCTACGCCAGCAGGTGGAGGCTTGTTCAGGGGTGGCGGCTTTTCAGCC 1871
Db 2121 GAGGCGAGAGGATCTACGCCAGCAGGTGGAGGCTTGTTCAGGGGTGGCGGCTTTTCAGCC 2180
QY 1872 CTCTGGCTTGGCTTGTGCTTTCACAGTG 1899
Db 2181 CTCTGGCTTGGCTTGTGCTTTCACAGTG 2208

RESULT 14
ADK19410
ID ADK19410 standard; cDNA; 2711 BP.
XX
AC ADK19410;
XX
DT 17-JUN-2004 (first entry)
XX
DE Partial cDNA encoding human LGR6 polypeptide #2.
XX
KW G-protein coupled receptor; GPCR; large G-protein coupled receptor 6;
LGR6; weight disorder; neural disorder; endocrine disorder;
KW cardiovascular disorder; anorectic; anabolic; immunomodulator;
KW neuroprotective; nootropic; antiparkinsonian; anti-HIV; anticonvulsant;
KW antihypertensive; osteopathic; cardiac; vasotropic; hypotensive;
KW antiarteriosclerotic; antiarrhythmic; human; gene; ss.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
CDS 1..2211
FT /*tag= a
FT /partial
FT /product= "LGR6 polypeptide #2"
FT /note= "This sequence lacks a start codon"
XX
PN US2004058377-A1.
XX
PD 25-MAR-2004.
XX
PF 18-SEP-2003; 2003US-00664667.
XX
PR 06-MAY-1999; 99US-0132896P.
PR 08-MAY-2000; 2000US-00566588.
PR 08-MAY-2001; 2001US-00851595.
XX
PA (MILL-) MILLENNIUM PHARM INC.
XX
PI Gu W;
XX
```

DR	WPI; 2004-268789/25.	Qy	612	GCACACCCAGTGTCCAGTGTAGCCCTACTCCAGGCCCTTCAAGCCCTGTGTAGTACCT	671
DX	P-PSDB; ADK19411.	Dd	921	GCACACCCAGTGTCCAGTGTAGCCCTACTCCAGGCCCTTCAAGCCCTGTGTAGTACCT	980
PT	New large G-protein coupled receptor 6 nucleic acid molecules and	Qy	672	CTTTGAAAGCTGGGGCATCCGCTTGGCGTGTGGGCAATCGTGTGTCTCCGCTGTCTG	731
PT	polypeptides, useful for diagnosing, preventing or treating diseases	Dd	981	CTTTGAAAGCTGGGGCATCCGCTTGGCGTGTGGGCAATCGTGTGTCTCCGCTGTCTG	1040
PT	associated with aberrant nucleic acid or protein activity, e.g. obesity,				
PT	anorexia or cachexia.				
XX					
PS	Claim 2; SEQ ID NO 7; 145pp; English.	Qy	732	CAATGGAATGGTGTCTGTACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	791
XX		Dd	1041	CAATGGAATGGTGTCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	1100
CC	The present invention relates to the isolation of novel members of the G-	Qy	792	GTTCGTGATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	851
CC	protein coupled receptor (GPCR) family designated as large G-protein	Dd	1101	GTTCGTGATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	1160
CC	coupled receptor 6 (LGR6), and the polynucleotide sequences encoding	Qy	852	AGCTTCAGTTCGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	911
CC	them. The invention also discloses LGR6 fusion proteins, antigenic	Dd	1161	AGCTTCAGTTCGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	1220
CC	peptides, and anti-LGR6 antibodies, recombinant expression vectors, host	Qy	912	GGGCTAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	971
CC	cells, and non-human transgenic animals in which an LGR6 gene has been	Dd	1221	GGGCTAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	1280
CC	introduced or disrupted. The LGR6 polynucleotide sequences, polypeptide	Qy	972	GCTGCTCAGTTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	1031
CC	sequences, and antibodies are useful for diagnosing, treating or	Dd	1281	GCTGCTCAGTTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	1340
CC	preventing diseases associated with aberrant LGR6 expression or activity,	Qy	1032	GAAGTTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	1091
CC	such as weight disorders (e.g. obesity, anorexia nervosa or cachexia),	Dd	1341	GAAGTTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	1400
CC	neural disorders (e.g. central nervous system disorder, including	Qy	1092	GCTGCTCAGTTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	1151
CC	Alzheimer's disease, Parkinson's disease, AIDS-related dementia, multiple	Dd	1401	GCTGCTCAGTTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	1460
CC	sclerosis and Huntington's disease), endocrine disorders (e.g.	Qy	1152	GCCCTAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	1211
CC	hypothyroidism, hyperthyroidism or acromegaly), or cardiovascular	Dd	1461	GCCCTAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	1520
CC	disorders (e.g. ischaemia-reperfusion injury, coronary artery disease,	Qy	1212	GATGAATCTCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	1271
CC	myocardial infarction, arrhythmia, atherosclerosis, hypertension or	Dd	1521	GATGAATCTCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	1580
CC	congestive heart failure). The LGR6 polynucleotide and polypeptide	Qy	1272	GCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	1331
CC	sequences are also useful as targets for developing modulating agents	Dd	1581	GCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	1640
CC	that regulate a variety of cellular processes, e.g. neural and endocrine	Qy	1332	CATCTTCGACAGCGGGCTCTCTACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	1391
CC	processes and thermogenesis. They can be used in screening assays (e.g.	Dd	1641	CATCTTCGACAGCGGGCTCTCTACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	1700
CC	chromosome mapping, tissue typing or in forensic biology), or in	Qy	1392	GGGCTCTCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	1451
CC	predictive medicine (e.g. diagnostic assays, prognostic assays,	Dd	1701	GGGCTCTCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	1760
CC	monitoring clinical trials or pharmacogenomics). The present sequence	Qy	1452	GCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	1511
CC	encodes a human LGR6 polypeptide.	Dd	1761	GCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	1820
XX		Qy	1512	TCGGCGCTTTCGGCGCCCGCGCAGGGGACTCAGGGCCCTTATGCTGCTGCGGCGGGGA	1571
SQ	Sequence 2711 BP; 530 A; 837 C; 740 G; 604 T; 0 U; 0 Other;	Dd	1821	TCGGCGCTTTCGGCGCCCGCGCAGGGGACTCAGGGCCCTTATGCTGCTGCGGCGGGGA	1880
		Qy	1572	GCTGGAGAAGAGCTCTCTGTGATTTCTACCCAGGGCCCTGGTAGCCCTTCTCTGATGTGGATCT	1631
		Dd	1881	GCTGGAGAAGAGCTCTCTGTGATTTCTACCCAGGGCCCTGGTAGCCCTTCTCTGATGTGGATCT	1940
		Qy	1632	CATTCTGGAAGCTTCTGGAAGCTGGGGCGCCCTTGGGCTGGAGACCTATGGCTTCCCTC	1691
		Dd	1941	CATTCTGGAAGCTTCTGGAAGCTGGGGCGCCCTTGGGCTGGAGACCTATGGCTTCCCTC	2000
		Qy	1692	AGTGACCTCATCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	1751

Query Match	92.4%;	Score	1754.4;	DB	12;	Length	2711;
Best Local Similarity	99.9%;	Pred. No.	0;				
Matches	1766;	Conservative	0;	Mismatches	1;	Indels	1;
Qy	132	AGCGGCTTCCAAACCAACCGCATCTGGGAAATTTGGAGCTGACACCTTCAGCCAGCTGAG	191				
Dd	442	ATCGGCTTCCAAACCAACCGCATCTGGGAAATTTGGAGCT-GACACCTTCAGCCAGCTGAG	500				
Qy	192	CTCCCTCGAAGCCCTTGGATCTTAGCTGGAAACCGCATCCGGTCCATCCACCTGAGGCTT	251				
Dd	501	CTCCCTCGAAGCCCTTGGATCTTAGCTGGAAACCGCATCCGGTCCATCCACCTGAGGCTT	560				
Qy	252	CTCCACCTGACCTCCCTGCTGCTGAGCTGAGCTGAGCAACACCGCTGACCACTGCC	311				
Dd	561	CTCCACCTGACCTCCCTGCTGCTGAGCTGAGCTGAGCAACACCGCTGACCACTGCC	620				
Qy	312	CCTGCTGAGCTTGGGGCTTGGTGAATCTGAAGCTCAAGGGAACCTTGTCTCTCCCA	371				
Dd	621	CCTGCTGAGCTTGGGGCTTGGTGAATCTGAAGCTCAAGGGAACCTTGTCTCTCCCA	680				
Qy	372	GGCTTCTTCCAAAGCAGTTTCCCAAACTGAGGATCTTGGAGTGGCTTATGCTTACCA	431				
Dd	681	GGCTTCTTCCAAAGCAGTTTCCCAAACTGAGGATCTTGGAGTGGCTTATGCTTACCA	740				
Qy	432	GTGCTGCTCCCTATGGGATGTGTGCGAGCTTCTCAAGGCTCTGGGAGTGGGAGGTGA	491				
Dd	741	GTGCTGCTCCCTATGGGATGTGTGCGAGCTTCTCAAGGCTCTGGGAGTGGGAGGTGA	800				
Qy	492	AGACTTTCACCTTGTATGATGAGGAGTCTTCAAAAGGCGCTTGGGCTCTTTCGCAGACA	551				
Dd	801	AGACTTTCACCTTGTATGATGAGGAGTCTTCAAAAGGCGCTTGGGCTCTTTCGCAGACA	860				
Qy	552	AGCAGAGAACCACTATGACCAAGGACCTTGGATGAGCTCCAGCTGGAGATGGAGGACTCAAA	611				
Dd	861	AGCAGAGAACCACTATGACCAAGGACCTTGGATGAGCTCCAGCTGGAGATGGAGGACTCAAA	920				

Db 2001 AGTGACCTCTATCTCTGTACAGCCAGGAGGCGCCCAAGGCTGGAGGAGCCATTTGCT 2060
Qy 1752 AGAGCCAGAGGGGAACCACTTTGGGAAACCCCAACCCCTCCATGGATGAGAACTCTGCT 1811
Db 2061 AGAGCCAGAGGGGAACCACTTTGGGAAACCCCAACCCCTCCATGGATGAGAACTCTGCT 2120
Qy 1812 GAGGCGAGAGGATCTAGCCAGCAGGTGGAGGCTTTGTCAGGGGTGCGGCTTTACGCC 1871
Db 2121 GAGGCGAGAGGATCTAGCCAGCAGGTGGAGGCTTTGTCAGGGGTGCGGCTTTACGCC 2180
Qy 1872 CTCTGGCTTGGCTTGTCTTACAGTG 1899
Db 2181 CTCTGGCTTGGCTTGTCTTACAGTG 2208
RESULT 15
ID AAI67927
XX ID AAI67927 standard; cDNA; 2901 BP.
AC AAI67927;
XX
DT 13-MAR-2002 (first entry)
XX
DE Human LGR6 polypeptide coding sequence (clone Fbh150881).
XX
KW Large G-protein coupled receptor 6; LGR6; GPCR; immunomodulator;
XX anti-HIV; cytosolic; neurotrophic; neuroprotective; antiparkinsonian;
KW anticonvulsant; neuroleptic; antidepressant; tranquilizer; antimigraine;
KW osteopathic; antitumor; vasodilator; vasopressor; hypotensive; cardiac;
KW antiarrhythmic; anorectic; gene therapy; human; ss.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 1..2901
FT /*tag= a
FT /product= "LGR6 polypeptide"
PN WO200185768-A2.
XX
PD 15-NOV-2001.
XX
PF 08-MAY-2001; 2001WO-US015002.
XX
PR 08-MAY-2000; 2000US-00566588.
XX
PA (MILL-) MILLENNIUM PHARM INC.
XX
PI Gu W;
XX
WPI, 2002-055584/07.
DR P-P8DB; AAG66141.
XX
PT Novel isolated large G-protein coupled receptor 6 polypeptide, useful for
PT diagnosing and treating weight disorder, metabolic disorder, central
PT nervous system disorder, endocrine disorder and cardiovascular disorder.
XX
PS Claim 2; Fig 14; 198pp; English.
XX
CC The invention provides isolated large G-protein coupled receptor 6 (LGR6)
CC polypeptides and polynucleotides encoding them. The LGR6 polypeptides and
CC polynucleotides are useful as targets for developing modulating agents
CC that regulate a variety of cellular processes, e.g. neural and endocrine
CC processes, as well as thermogenesis. They are useful for developing novel
CC diagnostic and therapeutic agents for LGR6 associated disorders such as
CC weight disorders (anorexia, obesity), eating disorders, acquired
CC immunodeficiency syndrome (AIDS)-related wasting, cancer-related wasting,
CC metabolic disorders, central nervous system disorders (e.g. Alzheimer's
CC disease, Parkinson's disease, Huntington's disease, Gilles de la
CC Tourette's syndrome, multiple sclerosis, amyotrophic lateral sclerosis,
CC epilepsy, Jakob-Creutzfeldt disease, depression, schizophrenia,
CC obsessive-compulsive disorder, anxiety, panic disorder, migraine),

CC endocrine disorders (hypo and hyperthyroidism, dwarfism, gigantism,
CC acromegaly), and cardiovascular disorders (atherosclerosis, restenosis,
CC hypertension, myocardial infarction, arrhythmia). The LGR6 polypeptides,
CC polynucleotides and antibodies are useful in screening assays, detection
CC assays (e.g. chromosomal mapping, tissue typing, forensic biology),
CC predictive medicine (e.g. diagnostic assays, prognostic assays,
CC monitoring clinical trials and pharmacogenomics), and in methods of
CC treatment (e.g. therapeutic and prophylactic). The present sequence
XX represents a human LGR6 polypeptide coding sequence
SQ Sequence 2901 BP; 513 A; 973 C; 818 G; 597 T; 0 U; 0 Other;
Query Match 92.4%; Score 1754.4; DB 6; Length 2901;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1766; Conservative 0; Mismatches 1; Indels 1; Gaps 1;
Qy 132 AGCGGCTCCAAACAACCGCATCTGGGAATTTGGAGTNGACACCTTCACCCAGCTGAG 191
Db 1135 ATCGGCTCCAAACAACCGCATCTGGGAATTTGGAGTNGACACCTTCACCCAGCTGAG 1193
Qy 192 CTCCTGCAAGCCCTGGATCTTAGCTGGAACGCCATCCGGTCCATCCACCTTGAGGCCCT 251
Db 1194 CTCCTGCAAGCCCTGGATCTTAGCTGGAACGCCATCCGGTCCATCCACCTTGAGGCCCT 1253
Qy 252 CTCACCCCTGCATCTCCCTGTCAGCTGGAACGCCATCCGGTCCATCCACCTTGAGGCCCT 311
Db 1254 CTCACCCCTGCATCTCCCTGTCAGCTGGAACGCCATCCGGTCCATCCACCTTGAGGCCCT 1313
Qy 312 CTTGGCTGGATCTGGGGGCTTGATGATCTGAACTCAAGAGGAACTTCTCTCTCCCA 371
Db 1314 CTTGGCTGGATCTGGGGGCTTGATGATCTGAACTCAAGAGGAACTTCTCTCTCCCA 1373
Qy 372 GGCCTTCTCAAGGACAGTTTCCCAAACTGAGGATCTTGAGGCTGCTTATGCTTACCA 431
Db 1374 GGCCTTCTCAAGGACAGTTTCCCAAACTGAGGATCTTGAGGCTGCTTATGCTTACCA 1433
Qy 432 GTGCTGCTCCATATGGGATGTGTCAGCTTCTTCAAGGCTCTGGGAGTGGGAGGCTGA 491
Db 1434 GTGCTGCTCCATATGGGATGTGTCAGCTTCTTCAAGGCTCTGGGAGTGGGAGGCTGA 1493
Qy 492 AGACCTTCACCTTGATGATGAGGATCTTCAAAAGGCGCCCTGGGCTCTTGGCAGACA 551
Db 1494 AGACCTTCACCTTGATGATGAGGATCTTCAAAAGGCGCCCTGGGCTCTTGGCAGACA 1553
Qy 552 AGCAGAGAACCACTATGACAGGACCTGGATGAGCTCCAGCTGGAGATGGAGGACTCAA 611
Db 1554 AGCAGAGAACCACTATGACAGGACCTGGATGAGCTCCAGCTGGAGATGGAGGACTCAA 1613
Qy 612 GCCACACCCAGTGTCCAGTGTAGCCCTACTCCAGGCGCCCTTCAAGGCTCTGAGTACCT 671
Db 1614 GCCACACCCAGTGTCCAGTGTAGCCCTACTCCAGGCGCCCTTCAAGGCTCTGAGTACCT 1673
Qy 672 CTTTGAAGCTGGGGCATCCGCTGGCGGTGGGCGCATCGTGTGCTCTCGTGTCTG 731
Db 1674 CTTTGAAGCTGGGGCATCCGCTGGCGGTGGGCGCATCGTGTGCTCTCGTGTCTG 1733
Qy 732 CAATGGACTGTGCTGTGCTGACCGGTTCGCTGGGGGCGCTGCGCCCTTCAAGGCTCAA 791
Db 1734 CAATGGACTGTGCTGTGCTGACCGGTTCGCTGGGGGCGCTGCGCCCTTCAAGGCTCAA 1793
Qy 792 GTTTGTGGTGGATGCGATTCGAGCGCGCAACACTTGAATGGCATTTCTGTGCGCTTCT 851
Db 1794 GTTTGTGGTGGATGCGATTCGAGCGCGCAACACTTGAATGGCATTTCTGTGCGCTTCT 1853
Qy 852 AGCTCTAGTCGATCCCTGACCTTTGCTGATTCAGTTCCTGAGTACGAGCGCGCTGGAGAC 911
Db 1854 AGCTCTAGTCGATCCCTGACCTTTGCTGATTCAGTTCCTGAGTACGAGCGCGCTGGAGAC 1913
Qy 912 GGGCTAGGCTGCGGGGCGCACTGCTTCTGGGAGTACTTGGGTGGAGGATCGGTGCT 971
Db 1914 GGGCTAGGCTGCGGGGCGCACTGCTTCTGGGAGTACTTGGGTGGAGGATCGGTGCT 1973
Qy 972 GCTGCTCACTCTGCGCGCAGTGCAGGCTCTCGCTCTCTGCTGTGTCGGGCTATGG 1031

Db	1974		GCTGCTCACTCTGGCCGAGTGCACTGAGCGTCTCCGCTCTCCGTGTCCTGGGCTATGG	2033
Qy	1032		GAAGTCCCCCTCCCTGGGCGAGGTTTCGAGCAGGGGTCTAGGCTGCCCTGGGCACTGGCAGG	1091
Db	2034		GAAGTCCCCCTCCCTGGGCGAGGTTTCGAGCAGGGGTCTAGGCTGCCCTGGGCACTGGCAGG	2093
Qy	1092		GCTGGCCGCGCACTGCGCCCTGGGCTCAGTGGGAGATACGGGGCTTCCCACTCTGCCT	1151
Db	2094		GCTGGCCGCGCACTGCGCCCTGGGCTCAGTGGGAGATACGGGGCTTCCCACTCTGCCT	2153
Qy	1152		GCCCTAGCGGCACTGAGGGTTCAGCCAGCAGCCCTGGGCTTCCCGTGGGCCCTGGTGAT	1211
Db	2154		GCCCTAGCGGCACTGAGGGTTCAGCCAGCAGCCCTGGGCTTCCCGTGGGCCCTGGTGAT	2213
Qy	1212		GATGAATCTCTTCTGTTTCTGGTGGTGGCCGGTGCCTACATCAAACTGTACTGTGACCT	1271
Db	2214		GATGAATCTCTTCTGTTTCTGGTGGTGGCCGGTGCCTACATCAAACTGTACTGTGACCT	2273
Qy	1272		GCCGCGGGGCACTTTGAGGCGGTGAGGCACTGGCCATGCTGAGGCACTGGGCTGGCT	1331
Db	2274		GCCGCGGGGCACTTTGAGGCGGTGAGGCACTGGCCATGCTGAGGCACTGGGCTGGCT	2333
Qy	1332		CATCTTGCAGACGGGCTCTCTACTGTCCGTTGGCTTCTCAGCTTTCGGCTCCATGCT	1391
Db	2334		CATCTTGCAGACGGGCTCTCTACTGTCCGTTGGCTTCTCAGCTTTCGGCTCCATGCT	2393
Qy	1392		GGGCTCTTCTCTGTACGCCCGAGGCGGTCAAGTCTGTCTGTCTGTGGTGGTGGCCCT	1451
Db	2394		GGGCTCTTCTCTGTGTACGCCCGAGGCGGTCAAGTCTGTCTGTCTGTGGTGGTGGCCCT	2453
Qy	1452		GCCTGCCTGCTCAACCACTGCTGTACTCTCTCAACCCCACTTCCGGGATGACCT	1511
Db	2454		GCCTGCCTGCTCAACCACTGCTGTACTCTCTCAACCCCACTTCCGGGATGACCT	2513
Qy	1512		TCGGCGGCTTCGGCCCGCGCAGGGGACTCAGGGCCCCCTAGCCTATGCTGGGCGGGGA	1571
Db	2514		TCGGCGGCTTCGGCCCGCGCAGGGGACTCAGGGCCCCCTAGCCTATGCTGGGCGGGGA	2573
Qy	1572		GCTGGAGAAGAGCTCTGTGATTCTACCCAGGCTCTGTAGCCTTCTCTGATGTGGATCT	1631
Db	2574		GCTGGAGAAGAGCTCTGTGATTCTACCCAGGCTCTGTAGCCTTCTCTGATGTGGATCT	2633
Qy	1632		CATTCTGGAGCTTCTGAGCTGGGCGGCCCTGGGCTGGAGACCTATGGCTTCCCTC	1691
Db	2634		CATTCTGGAGCTTCTGAGCTGGGCGGCCCTGGGCTGGAGACCTATGGCTTCCCTC	2693
Qy	1692		AGTGACCTCTATCTCTGTACAGCAGCCAGGGGCCCCCAGGCTGGAGGCGAGCCATTGTGT	1751
Db	2694		AGTGACCTCTATCTCTGTACAGCAGCCAGGGGCCCCCAGGCTGGAGGCGAGCCATTGTGT	2753
Qy	1752		AGAGCCAGAGGGGAACCACTTTTGGGAACCCCAACCTCTCCATGGATGGAGAACTGTGCT	1811
Db	2754		AGAGCCAGAGGGGAACCACTTTTGGGAACCCCAACCTCTCCATGGATGGAGAACTGTGCT	2813
Qy	1812		GAGGGCAGAGGATCTAGCCAGCAGGTGGAGGCTTCTCAGGGGCTGGCGGCTTTCAGCC	1871
Db	2814		GAGGGCAGAGGATCTAGCCAGCAGGTGGAGGCTTCTCAGGGGCTGGCGGCTTTCAGCC	2873
Qy	1872		CTCTGGCTTGGCTTTGCTTTCACACGTG	1899
Db	2874		CTCTGGCTTGGCTTTGCTTTCACACGTG	2901

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 14, 2005, 11:57:53 ; Search time 4096.82 Seconds
(without alignments)
17643.956 Million cell updates/sec

Title: US-10-664-667-6

Perfect score: 1899

Sequence: 1 aatacgaactactataggga.....tggcctttgttcacacgtg 1899

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 34239544 seqs, 19032134700 residues

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST:*

1: gb_est1:*

2: gb_est2:*

3: gb_hlc:*

4: gb_est3:*

5: gb_est4:*

6: gb_est5:*

7: gb_est6:*

8: gb_gse1:*

9: gb_gse2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1811.6	95.4	3509	3 BC038795	BC038795 Homo sapi
2	1325	69.8	2822	3 AK052873	AK052873 Mus muscu
3	1298.2	68.4	3583	3 AK085901	AK085901 Mus muscu
4	537	28.3	982	2 BF159363	BF159363 601770339
5	534.4	28.1	819	4 BG863804	BG863804 602798701
6	516.2	27.2	812	4 BG916782	BG916782 602816031
7	481.4	25.4	690	2 BB636845	BB636845 BB636845
8	479.8	25.3	3855	3 AK040883	AK040883 Mus muscu
9	465.8	24.5	680	7 CN258480	CN258480 170005321
10	457.6	24.1	753	4 BG916963	BG916963 602816251
11	424.4	22.3	869	2 BF158974	BF158974 601766476
12	385.2	20.3	461	2 BF149368	BF149368 RC1-HT025
13	362.4	19.1	2724	9 AV411734	AV411734 Mus muscu
14	355.6	18.7	2724	9 AV411732	AV411732 Homo sapi
15	322.4	17.0	640	7 CO351740	CO351740 DR_AOV_NR
16	282.4	14.9	310	7 Z44115	Z44115 HSCISG111 n
17	274	14.4	2593	9 AV411733	AV411733 Pan trogl
18	265.4	14.1	1719	2 BF144122	BF144122 601788812
19	256.4	13.5	471	5 BX281802	BX281802 BX281802
20	256.4	13.5	771	5 BQ425502	BQ425502 AGENCOURT
21	239.2	12.6	3006	3 AK033699	AK033699 Mus muscu
22	237.6	12.5	3035	3 AK044357	AK044357 Mus muscu
23	225.4	11.9	282	2 BF375245	BF375245 QVO-ST021
24	209.4	11.0	515	6 CA034787	CA034787 4000482 B

25	203	10.7	1212	4 BM543032	BM543032 AGENCOURT
26	200.6	10.6	671	4 BM795259	BM795259 K-EST0077
27	200.6	10.6	776	5 BU708439	BU708439 UI-M-FCO-
28	195	10.3	1104	1 AL530798	AL530798 AL530798
29	189.2	10.0	467	4 BM673325	BM673325 UI-B-CRI-
30	169.2	8.9	448	1 AI596321	AI596321 ve23f10.y
31	166.2	8.8	651	2 BB621262	BB621262 BB621262
32	166	8.7	800	4 BG196708	BG196708 RST15934
33	159.2	8.4	798	4 BI143698	BI143698 602907345
34	157.2	8.3	557	1 AA435175	AA435175 ve23f10.x
35	155.2	8.2	807	3 AK016357	AK016357 Mus muscu
36	155.2	8.2	816	6 BY716002	BY716002 BY716002
37	153.8	8.1	695	6 BY731701	BY731701 BY731701
38	150.2	7.9	607	5 BQ572012	BQ572012 UI-M-FB0-
39	138.2	7.3	880	5 BQ216343	BQ216343 AGENCOURT
40	126.4	6.7	181	7 Z38618	Z38618 HSC01D122 n
41	125.6	6.6	865	9 CNS022RR	AL221328 Tetraodon
42	124.8	6.6	839	4 BI150746	BI150746 602314834
43	122.8	6.5	386	6 CA035351	CA035351 4001250 B
44	122.2	6.4	494	7 CN683525	CN683525 E0180A02-
45	121.2	6.4	391	8 B53762	B53762 CIT-HSP-201

ALIGNMENTS

RESULT 1	BC038795	3509 bp	mRNA	linear	HTC 19-NOV-2003
LOCUS	BC038795				
DEFINITION	Homo sapiens leucine-rich repeat-containing G protein-coupled receptor 6, mRNA (cdna clone IMAGE:5220507), with apparent retained intron.				
ACCESSION	BC038795				
VERSION	BC038795.1				GI:24433474
KEYWORDS	HTC.				
SOURCE	Homo sapiens				
ORGANISM	Homo sapiens				
REFERENCE	1 (bases 1 to 3509)				
AUTHORS	Strausberg,R.L., Feingold,E.A., Grouse,L.H., Derge,J.G., Klausner,R.D., Collins,F.S., Wagner,L., Shenmen,C.M., Schuler,G.D., Altschul,S.F., Zeeberg,B., Buettow,K.H., Schaefer,C.F., Bhat,N.K., Hopkins,R.F., Jordan,H., Moore,T., Max,S.I., Wang,J., Haleh,F., Diatchenko,L., Marusina,K., Farmer,A.A., Rubin,G.M., Hong,L., Stapleton,M., Soares,M.B., Bonaldo,M.F., Casavant,T.L., Scheetz,T.E., Brownstein,M.J., Usdin,T.B., Toshiyuki,S., Carninci,P., Prange,C., Raha,S.S., Loquellano,N.A., Peters,G.J., Abramson,R.D., Mullaly,S.J., Bosak,S.A., McEwan,P.J., McKernan,K.J., Malek,J.A., Gunaratne,P.H., Richards,S., Worley,K.C., Hale,S., Garcia,A.M., Gay,L.J., Hulyk,S.W., Villalon,D.K., Muzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A., Fahey,J., Helton,E., Kettman,M., Madan,A., Rodriguez,S., Sanchez,A., Whiting,M., Madan,A., Young,A.C., Shevchenko,Y., Bouffard,G.G., Blakesley,R.W., Touchman,J.W., Green,E.D., Dickson,M.C., Rodriguez,A.C., Grimwood,J., Schmutz,J., Myers,R.M., Butlerfield,Y.S., Krzywinski,M.I., Skalka,U., Smalhus,D.E., Schnerch,A., Schein,J.E., Jones,S.J. and Marra,M.A.				
TITLE	Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences				
JOURNAL	Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)				
MEDLINE	22388257				
PUBMED	12477932				
REFERENCE	2 (bases 1 to 3509)				
AUTHORS	Strausberg,R.				
TITLE	Direct Submission				
JOURNAL	Submitted (25-OCT-2002) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA				
REMARK	NIH-MGC Project URL: http://mgc.nci.nih.gov				
COMMENT	Contact: MGC help desk Email: cgapbs-r@mail.nih.gov				

Tissue Procurement: Life Technologies, Inc.
 cDNA Library Preparation: Life Technologies, Inc.
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: National Institutes of Health Intramural
 Sequencing Center (NISC),
 Gaithersburg, Maryland;
 Web site: <http://www.nisc.nih.gov/>
 Contact: nisc.mc@nih.gov
 Akhter, N., Ayle, K., Beckstrom-Sternberg, S.M., Benjamin, B.,
 Blakesley, R.W., Bouffard, G.G., Breen, K., Brinkley, C., Brooks, S.,
 Dietrich, N.L., Granite, S., Guan, X., Gupta, J., Haghighi, P.,
 Hansen, N., Ho, S.-L., Karlins, E., Kwong, P., Lalic, P., Legaspi, R.,
 Maduro, O.L., Masiello, C., Maskeri, B., Mastrian, S.D., McCloskey, J.C.,
 McDowell, J., Pearson, R., Stantripop, S., Thomas, P.J., Touchman, J.W.,
 Tsurgoun, C., Vogt, J.L., Walker, M.A., Wetherby, K.D., Wiggins, L.,
 Young, A., Zhang, L.-H. and Green, E.D.

Clone distribution: MGC clone distribution information can be found
 through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
 Series: IRAC Plate: 64 Row: n Column: 17
 This clone was selected for full length sequencing because it
 passed the following selection criteria: Hexamer frequency ORF
 analysis

This clone has the following problem: retained intron.

FEATURES

Location/Qualifiers
 1..3509
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:5220507"
 /tissue type="Pancreas, Spleen, adult pooled"
 /clone_lib="NTH MGC_120"
 /lab_host="DH10B"
 /note="Vector: pCMV-SPORT6"

ORIGIN

Query Match 95.4%; Score 1811.6; DB 3; Length 3509;
 Best Local Similarity 99.7%; Pred. No. 0;
 Matches 1825; Conservative 0; Mismatches 4; Indels 1; Gaps 1;
 70 GCGTCCTGGAGCGAGCAGGGTCTGAGCTGCGGCTCATCCAGGCTCTCTTGTGCTGCC 129
 1063 GGGGGCATGGAGCGAGCGAGGGTCTGAGCTGCGGCTCATCCAGGCTCTCTTGTGCTGCC 1122
 130 CTAGCGGCTCCACACACCGCATCTGGGAAATGGAGCTNGACACCTTCAGCCAGCTG 189
 1123 CTAGCGGCTCCACACACCGCATCTGGGAAATGGAGCT-GACACCTTCAGCCAGCTG 1181
 190 AGCTCCCTGCAAGCCCTGGATCTTAGCTGGAAAGCCATCCGCTCCATCCACCTTGAGGCC 249
 1182 AGCTCCCTGCAAGCCCTGGATCTTAGCTGGAAAGCCATCCGCTCCATCCACCTTGAGGCC 1241
 250 TTCTCCACCTTGCACTCCCTGGTCAAGCTGAGCTGACCTGACAGAACCCAGCTGACCACTG 309
 1242 TTCTCCACCTTGCACTCCCTGGTCAAGCTGAGCTGACCTGACAGAACCCAGCTGACCACTG 1301
 310 CCCCTGGGTGACCTTGGGGGCTTGATGCACTGAAGCTCAAGGGAACCTTGCTCTCTCC 369
 1302 CCCCTGGGTGACCTTGGGGGCTTGATGCACTGAAGCTCAAGGGAACCTTGCTCTCTCC 1361
 370 CAGGCTCTTCCAAAGGACAGTTTCCCAAACTGAGGATCCTGGAGGTGCTTATGCTCTAC 429
 1362 CAGGCTCTTCCAAAGGACAGTTTCCCAAACTGAGGATCCTGGAGGTGCTTATGCTCTAC 1421
 430 CAGTGTCTCCCTATGGGATGTGTGCCAGCTTCTTCAAGGCTCTGGGCAAGTGGAGGCT 489
 1422 CAGTGTCTCCCTATGGGATGTGTGCCAGCTTCTTCAAGGCTCTGGGCAAGTGGAGGCT 1481
 490 GAAGACCTTCACCTTGATGATGAGAGTCTTCAAAAAGGCCCTTCGGGCTCTCTGCGAGA 549
 1482 GAAGACCTTCACCTTGATGATGAGAGTCTTCAAAAAGGCCCTTCGGGCTCTCTGCGAGA 1541
 550 CAAGCAGAGAACCCTATGACAGGACCTTGGATGAGCTCCAGCTGGAGATGGAGGACTCA 609

1542 CAAGCAGAGAACCCTATGACAGGACCTGGATGAGCTCCAGCTGGAGATGGAGGACTCA 1601
 610 AAGCCACACCCAGTGTCCAGTGTAGCCCTACTCCAGGCCCTTCAAGCCCTGTGAGTAC 669
 1602 AAGCCACACCCAGTGTCCAGTGTAGCCCTACTCCAGGCCCTTCAAGCCCTGTGAGTAC 1661
 670 CTCTTTGAAAGCTGGGGCATCCGCTGCGCGTGTGGGCCATCGGTGCTCTCCGTGCTC 729
 1662 CTCTTTGAAAGCTGGGGCATCCGCTGCGCGTGTGGGCCATCGGTGCTCTCCGTGCTC 1721
 730 TGCAATGCACTGTGTGCTGACCGCTGTTCGCTGGCGGGCTGCCCCCTCCCGGCTG 789
 1722 TGCAATGCACTGTGTGCTGACCGCTGTTCGCTGGCGGGCTGCCCCCTGCCCCGCTG 1781
 790 AAGTTTGTGTAGGTGCGATTGAGGCGCAACACCTTGACTGCGCATTTCTGTGGGCTT 849
 1782 AAGTTTGTGTAGGTGCGATTGAGGCGCAACACCTTGACTGCGCATTTCTGTGGGCTT 1841
 850 CTAGCCTCAGTCCGATGCCCTGACCTTTGCTGCTGCTGAGTACGAGCCCGCTGGAG 909
 1842 CTAGCCTCAGTCCGATGCCCTGACCTTTGCTGCTGCTGAGTACGAGCCCGCTGGAG 1901
 910 ACGGGCTAGGCTGCCCGGCCACTTGGCTTCTTGGCAGTACTTGGGTGGAGGCACTCGGTG 969
 1902 ACGGGCTAGGCTGCCCGGCCACTTGGCTTCTTGGCAGTACTTGGGTGGAGGCACTCGGTG 1961
 970 CTGCTGTCTACTCTTGGCCGAGTGCAGTGCAGCGTCTCCGCTCTCTGTGCCGGCTAT 1039
 1962 CTGCTGTCTACTCTTGGCCGAGTGCAGTGCAGCGTCTCCGCTCTCTGTGCCGGCTAT 2021
 1030 GGGAGTCCCTCCCTGGCGAGGTTTCGAGCAGGGGCTTACGCTGCTGCTGGCACTGGCA 1089
 2022 GGGAGTCCCTCCCTGGCGAGGTTTCGAGCAGGGGCTTACGCTGCTGCTGGCACTGGCA 2081
 1090 GGGTGGCCCGCCACTGCGCCCTGGGCTCAGTGGGAGAAATACGGGGCTTCCCCACTCTGC 1149
 2082 GGGTGGCCCGCCACTGCGCCCTGGGCTCAGTGGGAGAAATACGGGGCTTCCCCACTCTGC 2141
 1150 CTGCTGTCTACTCTGAGGCTGAGCAGAGCCTCGGGCTTTCACCGTGGCCCTGGTG 1209
 2142 CTGCTGTCTACTCTGAGGCTGAGCAGAGCCTCGGGCTTTCACCGTGGCCCTGGTG 2201
 1210 ATGATGAACCTCTCTGTTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1269
 2202 ATGATGAACCTCTCTGTTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2261
 1270 CTGCGCGGGCGCACTTTGAGGCGCTGTTGGGACTGCGCCATGGTGGGCGACGTGGCCTGG 1329
 2262 CTGCGCGGGCGCACTTTGAGGCGCTGTTGGGACTGCGCCATGGTGGGCGACGTGGCCTGG 2321
 1330 CTGATCTTCGAGAGGGCTCTCTACTGTCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1389
 2322 CTGATCTTCGAGAGGGCTCTCTACTGTCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2381
 1390 CTGGGCTCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1449
 2382 CTGGGCTCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2441
 1450 CTGCTGCTCTCTCAACCCACTGCTGCTGCTGCTTCAACCCCACTTTCGGGGATGAC 1509
 2442 CTGCTGCTCTCTCAACCCACTGCTGCTGCTTCAACCCCACTTTCGGGGATGAC 2501
 1510 CTTTCGGCGCTTCGGGCCCGCGAGGGGACTCAGGGGCCCTAGCTATGCTGCGGCGGG 1569
 2502 CTTTCGGCGCTTCGGGCCCGCGAGGGGACTCAGGGGCCCTAGCTATGCTGCGGCGGG 2561
 1570 GAGCTGGAGAGAGCTCTGCTGATTTCTACCGAGCCCTGCTGCTGCTGCTGCTGCTGCTGCTG 1629
 2562 GAGCTGGAGAGAGCTCTGCTGATTTCTACCGAGCCCTGCTGCTGCTGCTGCTGCTGCTGCTG 2621
 1630 CTGATCTGGAAGCTTCTGAAGCTGGGCGGCCCTCGGGCTGGAGACTTATGGCTTCCCC 1689

Db	2622	CTCATTTCTGGAAGCTTCTGAAGCTGGGCGGCCCCCTGGGCTGGAGACTTATGGCTTCCCC	2681
Qy	1690	TCAGTGACCCCTCATCTCTCTGTGTCAGCAGCCAGGCGGCCCCAGGCTGGAGGGCAGCCATTGT	1749
Db	2682	TCAGTGACCCCTCATCTCTCTGTGTCAGCAGCCAGGCGGCCCCAGGCTGGAGGGCAGCCATTGT	2741
Qy	1750	GTAGAGCCAGAGGGGAACACATTTTGGGAACCCCAACCCCTCATGGATGGAGAACTGCTG	1809
Db	2742	GTAGAGCCAGAGGGGAACACATTTTGGGAACCCCAACCCCTCATGGATGGAGAACTGCTG	2801
Qy	1810	CTGAGGCGCAGAGGGATCTACGCCAGCAGTGGAGGCTTGTTCAGGGGGTGGCGGCTTTCAG	1869
Db	2802	CTGAGGCGCAGAGGGATCTACGCCAGCAGTGGAGGCTTGTTCAGGGGGTGGCGGCTTTCAG	2861
Qy	1870	CCCTCTGGCTTGGCTTTTCTTTCACACGTG	1899
Db	2862	CCCTCTGGCTTGGCTTTTCTTTCACACGTG	2891
RESULT 2			
LOCUS	AK052873	2822 bp	mRNA
DEFINITION	Mus musculus 16 days neonate heart cDNA, RIKEN full-length enriched library, clone:D830015D13 product:CDNA FLJ14471 FIS, CLONE MAMMA1001030, WEAKLY SIMILAR TO LUTROPIN-CHORIOGONADOTROPIC HORMONE RECEPTOR homolog [Homo sapiens], full insert sequence.		
ACCESSION	AK052873		
VERSION	AK052873.1	GI:26095426	
KEYWORDS	HTC; CAP trapper.		
SOURCE	Mus musculus (house mouse)		
ORGANISM	Mus musculus		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.		
AUTHORS	1 Carninci, P. and Hayashizaki, Y.		
TITLE	High-efficiency full-length cDNA cloning		
JOURNAL	Meth. Enzymol. 303, 19-44 (1999)		
MEDLINE	99279253		
PUBMED	10349636		
REFERENCE	2		
AUTHORS	Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.		
TITLE	Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes		
JOURNAL	Genome Res. 10 (10), 1617-1630 (2000)		
MEDLINE	20495374		
PUBMED	11042159		
REFERENCE	3		
AUTHORS	Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P., Konno, H., Akiyama, J., Nishi, K., Kitsunai, T., Tashiro, H., Itoh, M., Sumi, N., Iehi, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A., Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K., Fujiwaka, S., Inoue, K., Ozawa, Y., Izawa, M., Ohara, E., Wataniki, M., Yoneda, Y., Ishikawa, T., Tanaka, T., Matsura, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.		
TITLE	RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer		
JOURNAL	Genome Res. 10 (11), 1757-1771 (2000)		
MEDLINE	20530913		
PUBMED	11076861		
REFERENCE	4		
AUTHORS	The RIKEN Genome Exploration Research Group Phase II Team and the FANTOM Consortium.		
TITLE	Functional annotation of a full-length mouse cDNA collection		
JOURNAL	Nature 409, 685-690 (2001)		
REFERENCE	5		
AUTHORS	The FANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team.		
TITLE	Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs		
JOURNAL	Nature 420, 563-573 (2002)		
REFERENCE	6 (bases 1 to 2822)		
AUTHORS	Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Bono, H., Carninci, P., Fukuda, S., Furuno, M., Hanagaki, T., Hara, A., Hashizume, W., Hayashida, K., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hirozane, T., Hori, F., Imotani, K., Iehi, Y., Itoh, M., Kagawa, I., Kasukawa, T., Kato, H., Kawai, J., Kojima, Y., Kondo, S., Konno, H., Kouda, M., Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M., Nakamura, M., Nishi, K., Nomura, K., Numazaki, R., Ono, M., Ohsato, N., Okazaki, Y., Saito, R., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N., Sano, H., Sasaki, D., Shibata, K., Shinagawa, A., Shiraki, T., Sogabe, Y., Tagami, M., Tagawa, A., Takahashi, F., Takaku-Akahira, S., Takeda, Y., Tanaka, T., Tomaru, A., Toya, T., Yasunishi, A., Muramatsu, M. and Hayashizaki, Y.		
Direct Submission			
Submitted (16-JUL-2001)	Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Teurumi-ku, Yokohama, Kanagawa 230-0045, Japan [E-mail: genome-res@gsc.riken.jp, URL: http://genome.gsc.riken.jp/, Tel: 81-45-503-9222, Fax: 81-45-503-9216]		
COMMENT	CDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.		
	Please visit our web site for further details.		
URL: http://genome.gsc.riken.jp/			
URL: http://fantom.gsc.riken.jp/			
FEATURES	Location/Qualifiers		
source	1..2822		
	/organism="Mus musculus"		
	/mol_type="mRNA"		
	/strain="C57BL/6J"		
	/db_xref="FANTOM,DB:D830015D13"		
	/db_xref="taxon:10090"		
	/clone="D830015D13"		
	/tissue_type="heart"		
	/clone_lib="RIKEN full-length enriched mouse cDNA library"		
	/dev_stage="16 days neonate"		
misc_feature	1..2822		
	/note="CDNA FLJ14471 FIS, CLONE MAMMA1001030, WEAKLY SIMILAR TO LUTROPIN-CHORIOGONADOTROPIC HORMONE RECEPTOR homolog [Homo sapiens] (SPR[096K69, evidence: FASTY, 85.8%ID, 92.1%length, match=637])"		
ORIGIN			
Query Match	69.8%;	Score 1325;	DB 3: Length 2822;
Best Local Similarity	85.1%;	Pred. No. 7.3e-289;	
Matches 1503;	Conservative	0; Mismatches 261;	Indels 2; Gaps 2;
QY	135	GGCTTCCCAACCAACCGCATCTGGGAAATGGAGCTTGACACCTTCAGCCAGCTGAGCTC	194
DB	474	GGCTTCCCAACCAACCGCATCTGGGAAATGGAGCTTGACACCTTCAGCCAGCTGAGCTC	532
QY	195	CCTGCAAGCCCTGGATCTTACTGGACGCCATCCGCTCCATCCCTGAGGCTTCTC	254
DB	533	CTTGAAGCTTTAGACCTGAGTGGAAATGCCATCCGCTGCCATCCACCTGAGGCTTCTC	592
QY	255	CACCTGCACTCCCTGGTCAAGCTGGACCTGACAGCAACACGCTGACACACTGCCCT	314
DB	593	AACCTTCCGATCTTGGTTAAGCTGGACCTGACCTGACCAACGCTGACACACTGCCCT	652
QY	315	GGCTGGATCTGGGGCTTGGATGCATCTGAAGCTCAAGGGAACCTTCTCTCTCCAGGC	374
DB	653	GGCTGGCTGGAGGCTGATGCACCTGAGCTCAAGGGAACCTTGGCCCTGTCTCAGGC	712
QY	375	CTTCTCCAAGCAGTAGTTTCCAAAACCTGAGATCCTGGAGGTGCTTATGCCCTACCA	434
DB	713	CTTCTCCAAGGACAGTTCCTCCAAAACCTGAGATCCTGGAGGTGCTTATGCCCTAC	772
QY	435	CTGTCCCTTATGGATGTGTGCCAGCTTCTTCAGGCCCTCTGGCAGCTGGGAGCTG	494
DB	773	CTGTCCCTTACCGCATCTGTGCCAGCTTCTTCAGAGCCTCTGGGCGAGTGGCAGCC	832

QY	495	CCTTACCTTGATGATGATGAGGAGTCTTTCAAAAAGGCCCTCTGGGCTCTCTTGGCAGACAGC	554
DB	833	TTTTTCTATCAGAGGAAGAGGAGGACCAAAAGAGGCCCTCTGGTCTCTCTTGTCTGGACAAAGC	892
QY	555	AGAGAAACACTATGATACACAGGACCTGGATAGCTCCAGCTGCAGATGGAGACTCAAGACC	614
DB	893	TGAGAACCACTATGACCTTAGACCTTGGATGAGCTCCAGATGGGACACAGAGACTCAAGACC	952
QY	615	ACACCCAGTGTCCAGTGTAGCCCTTACTCCAGGCCCTTCAAGGCCCTGTGAGTAGCTCTTT	674
DB	953	ACACCCAGTGTCCAGTGTAGCCCTTACTCCAGGCCCTTCAAGGCCCTGTGAGTAGCTCTTT	1012
QY	675	TGAAAGCTGGGGCATCCGCCCTGGGCGGTGTGGCCCATGCTGTGCTCTCCGTGCTCGAA	734
DB	1013	TGAGAGCTGGGGCATCCGCCCTTGCTGTGTGGGCCCATCGTGTACTCTCCGTACTCTGTAA	1072
QY	735	TGGACTGTGTCTGTGACCGGTTCGCTGGGGGCGCTGCCCCCTGCCCCCGGTCAAGTT	794
DB	1073	CGGGCTGTGTCTGTGACAGTCTTTGCCAGCGGACCCAGGCCGCTGTCCCCCGGTCAAGCT	1133
QY	795	TGTGTAGGTTCGATTTGACAGCGGCCAACACCTTGACTTGGCATTTCTGTGGCCTTCTAGC	854
DB	1133	TGTGGTGGGTTCGATGGAGCGGCCAACGCCCTGTGGGCAATTTCTGTGTCTCTCTGGC	1192
QY	855	CTCAGTCGATCCCTCGAACCTTTTGTTCAGTTCCTTAGTACGAGGACCCGCTGGGAGACGG	914
DB	1193	CTCGTGGACGCTTGACTATGTTCAGTTTCGCTGAGTATGGAGCCCGCTGGGAGAGCG	1252
QY	915	GCTAGGCTGCCGGGCCACTGCTTCTTGGCAGTACTTGGGTTCGAGGACATCGTGTCTGCT	974
DB	1253	TCTGGGCTGCCAGGCTACGGGCTTCTTGGCTGTCTCTGGGTTTCAGAGCGCTGGGTGTCT	1312
QY	975	GCTCACTCTGGCCGACGTGACAGTTCAGCGTCTCTCGTCTCTCTGTCTCGGGCTATGGAA	1034
DB	1313	GCTCACACTGGCGGCGGTGACGTGCAGCATCTCCGTGACCTGCGTCCGAGCCTACGGGAA	1372
QY	1035	GTCCCCCTCTCTGGGACAGCTTCAGACAGAGGGTCTTAGGCTGCTGBCACTGGCAGGGCT	1094
DB	1373	GGCGCGTCTGCTGGCAGCGTCCGCGACGGCGCACTGGGATGCTGCGCTGGCGCGGCT	1432
QY	1095	GGCGCGGCACTGCCCTTGGGCTCAGTGGGGAGATACGGGGCTTCCGCACTCTGCTGCTGC	1154
DB	1433	GGCGCGACACTGCCCTTGGGCTCTGGTGGGAGATATGGCGGCTTCCGCACTCTGCTGCTGC	1492
QY	1155	CTACGCGCACCTCAGGGGTACGCAGACAGCCCTGGGCTTACCGTGGCCCTGGTGTGATGAT	1214
DB	1493	CTACGCGCCACCGAGGCGCGGCGGTGCTGGCTTCTGTGACCTCTGGTGTGATGAT	1552
QY	1215	GAACCTCTCTTGTCTCTGGTTCGTGGCGGCTGCCTACATCAAACTGCTGCTGACCTGCC	1274
DB	1553	GAACCTGCTCTGCTTCTGTGTGTGGCGGCGCTACATCAAGCTCTACTGTGACCTGCC	1612
QY	1275	GGGGGGGCACTTTGAGGCGGTGTGGGACTTGGCCCATGGTAGGCACTG - GGCTGTGGTCA	1333
DB	1613	ACGGGGTGACTTTGAGGCGGTGTGGGACTTGGCCCATGGTGGCCACGCTGGGCGCTGGCTCA	1672
QY	1334	TCTTTCGAGACGGGCTCTCTACTGTCCGTTGGGCTTCTCAGCTTCGCTCCCATGCTGG	1393
DB	1673	TCTTTGAGATGGGCTCTCTACTTGGCCCCGGTGGCTTCTCAGCTTTGGCTCTATGCTGG	1732
QY	1394	GCCTCTTCCCTGTCAAGCGGCGGCTCAAGTCTGTCTCTGGTGGTGTCTGCCCCCTGC	1453
DB	1733	GCCTCTTCCCTGTCAAGCGGCGGCTCAAGTCTGTCTCTGGTGGTGTCTGCTGCTGCTGC	1792
QY	1454	CTGCTCGCTCAACCCACTGTGTACTGTCTTTCAAACCCCACTTCGCGGATGACCTTC	1513
DB	1793	CTGCTCGCTCAACCCACTGTGTACTGTCTTTCAAACCCCACTTCGCGGATGACCTTC	1852
QY	1514	GGCGGCTTCCGCCCCGGCAGGAGACTCAGGGCCCCCTAGCCTATGCTTGGCGGCGGGAGC	1573
DB	1853	GGCGGCTTCTGGCCAAAGCCCTCGGTCCCAAGGCCCCCTAGCCTATGCTTGGAGCGGTGAGC	1912
QY	1574	TGGAGAAAGAGCTCTCTGTGATTTCTACCCAGGCGGCTGTGTAGGCTCTCTCTGATGTGATCTCA	1633

Db	1913	TGGAGAGAGCTCTCGACACTCCACCAAGCGCTGGTGGCTTTCTCAGATGTGGATCTTA	1972
Qy	1634	TTCTGGGAAGCTTCTGAAGCTGGGGCGGCCCTCGGCTGGAGACCTATGGCTTCCCTCCAG	1693
Db	1973	TTCTGGGAAGCTTCTGAAGCTGGGGCGGCCCTCGGCTGGAGACCTATGGCTTCCCTCCAG	2032
Qy	1694	TGACCTCATCTCTCTGTCAGCAGCAGCGGGGCCCCCCAGGCTGGAGGGCAGGCATTTGTGTAG	1753
Db	2033	TGACCTCATCTCTCGACATCAGCGGGGGCTACCAAGCTGGAGGAAACCAATTTCTAG	2092
Qy	1754	AGCCAGAGGGGAACCACTTTGGGAACCCCCAACCCCTCATGATGGAGAACTGCTGCTGA	1813
Db	2093	AGTCTGATGAAACCAAGTTTGGGAACCCCAACCTCCCATGAAGGGGAGAACTGCTGCTGA	2152
Qy	1814	GGGCAGAGGGATCTAGCCACAGCAGGTGGAGGCTTGTTCAGGGGGTGGCGGCTTTTCAGGCCCT	1873
Db	2153	AGGCAGAGGGAGGCCACTTTGGCAGGCTGTGGCTCTCCGTGGGTGGAGGCCCTCTGGGCCCT	2212
Qy	1874	CTGGCTTGGGCTTTGCTTTCACAGGTG	1899
Db	2213	CTGGCTCTCTTTGGCTCTCACTTG	2238
RESULT 3			
LOCUS	AK085901	3583 bp	linear
DEFINITION	Mus musculus 16 days neonate heart cDNA, RIKEN full-length enriched library, clone: D830026M09 product: CDNA FLJ14471 FIS, CLONE MAMMA1001030, WEAKLY SIMILAR TO LUTROPIN-CHORIOGNADOTROPIC HORMONE RECEPTOR homolog [Homo sapiens], full insert sequence.		
ACCESSION	AK085901		
VERSION	AK085901.1	GI:26103061	
KEYWORDS	HTC; CAP trapper.		
SOURCE	Mus musculus		
ORGANISM	Mus musculus (house mouse)		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.		
AUTHORS	1		
TITLE	Carninci, P. and Hayashizaki, Y.		
JOURNAL	High-efficiency full-length cDNA cloning		
MEDLINE	Meth. Enzymol. 303, 19-44 (1999)		
PUBMED	99279253		
REFERENCE	2		
AUTHORS	Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.		
TITLE	Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes		
JOURNAL	Genome Res. 10 (10), 1617-1630 (2000)		
MEDLINE	20499374		
PUBMED	11042159		
REFERENCE	3		
AUTHORS	Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P., Konno, H., Akiyama, J., Nishi, K., Kitesunai, T., Tashiro, H., Itoh, M., Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A., Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K., Fujiwaka, S., Inoue, K., Iogawa, Y., Izawa, M., Ohara, E., Wathiki, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsura, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.		
TITLE	RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer		
JOURNAL	Genome Res. 10 (11), 1757-1771 (2000)		
MEDLINE	20530913		
PUBMED	11076861		
REFERENCE	4		
AUTHORS	THE RIKEN Genome Exploration Research Group Phase II Team and the FANTOM Consortium.		
TITLE	Functional annotation of a full-length mouse cDNA collection		
JOURNAL	Nature 409, 685-690 (2001)		
REFERENCE	5		
AUTHORS	THE FANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team.		

TITLE Analysis of the mouse transcriptome based on functional annotation
JOURNAL Of 60,770 full-length cDNAs
REFERENCE Nature 420, 563-573 (2002)
AUTHORS 6 (bases 1 to 3583)

Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Bono, H., Carninci, P., Fukuda, S., Furuno, M., Hanezaki, T., Hara, A., Hashizume, W., Hayashida, K., Hayatsu, N., Hiramoto, K., Hiraoka, I., Hirozane, T., Hori, F., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kasukawa, T., Katoh, H., Kawai, J., Kojima, Y., Kondo, S., Konno, H., Kouda, M., Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M., Nakamura, M., Nishi, K., Nomura, K., Numazaki, R., Ohno, M., Ohsato, N., Okazaki, Y., Saito, R., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N., Sano, H., Sasaki, D., Shibata, K., Shinagawa, A., Shiraki, T., Sogabe, Y., Tgami, M., Tagawa, A., Takahashi, F., Takaku-Akahira, S., Takeda, Y., Tanaka, T., Tomaru, A., Toya, T., Yasunishi, A., Muramatsu, M. and Hayashizaki, Y.

TITLE Direct Submission
JOURNAL Submitted (16-APR-2002) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail: genome-res@sc.riken.jp, URL: http://genome.gsc.riken.jp/, Tel: 81-45-503-9222, Fax: 81-45-503-9216)

COMMENT cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.
 Please visit our web site for further details.
 URL: http://genome.gsc.riken.jp/
 URL: http://fantom.gsc.riken.jp/
 Location/Qualifiers

FEATURES
source
 1. 3583
 /organism="Mus musculus"
 /mol_type="mRNA"
 /strain="C57BL/6J"
 /db_xref="FANTOM_DB:D830026M09"
 /db_xref="taxon:10090"
 /clone="D830026M09"
 /tissue type="heart"
 /clone_lib="RIKEN full-length enriched mouse cDNA library"
 /dev_stage="16 days neonate"
 95. 3002
 /note="CDNA FLJ14471 FIS, CLONE MAMMAL001030, WEAKLY SIMILAR TO LUTROPIN-CHORIOGONADOTROPIC HORMONE RECEPTOR homolog [Homo sapiens] (SPTF|Q96K69, evidence: PASTY, 85.8%ID, 92.1%length, match=637)
 putative"
 3563. 3568
 /note="putative"
 3583
 /note="putative"

misc_feature
 95. 3002
 /note="CDNA FLJ14471 FIS, CLONE MAMMAL001030, WEAKLY SIMILAR TO LUTROPIN-CHORIOGONADOTROPIC HORMONE RECEPTOR homolog [Homo sapiens] (SPTF|Q96K69, evidence: PASTY, 85.8%ID, 92.1%length, match=637)
 putative"
 3563. 3568
 /note="putative"
 3583
 /note="putative"

polyA_signal
 3563. 3568
 /note="putative"
 3583
 /note="putative"

ORIGIN
 Query Match 68.4%; Score 1298.2; DB 3; Length 3583;
 Best Local Similarity 84.8%; Pred. No. 8.8e-283;
 Matches 1500; Conservative 0; Mismatches 264; Indels 4; Gaps 4;
 Qy 135 GGCCTCCACACACCGCATCGGAAATTGGAGCTNGACACCTTCAGCCAGCTGAGCTC 194
 Db 1233 GGCCTCCACATAACCGGATCAAGGAAATTGGTGC-AGATACCTTCAGCCAGCTGGGCTC 1291
 Qy 195 CCTCAAGCCCTGGATCTTACCTGGAAGCGCATCGGTCCATCCACCTCAGGCGCTTCTC 254
 Db 1292 CTTCGAAGCTTTAGACCTTAGTGGATGACCATCGTGCCATCCACCTCAGGCTTCTC 1351
 Qy 255 CACCTCGACCTCCCTGGTCAAGCTGGACCTGACAGACAAACAGCTGACACACTGCCCTC 314
 Db 1352 AACCTTCGATCTTGGTTAAGCTGGACCTGACTGACAAACAGCTGACCACTGCCCT 1411
 Qy 315 GGCCTGGACTTGGGGCTTGATGCATCTGAAGCTCAAAGGAAACCTTGCTCTCTCCAGGC 374

Db 1412 GGTGGGCTGGAGGCGCTGATGCACTGAAGCTCAAAGGAACTTGGGCCCTGTCTCAGGC 1471
 Qy 375 CTTCTCCAAAGACAGTTTTCCAAAACCTGAGGATCTCTGAGGTGCTTATGCTTACAGTG 434
 Db 1472 CTTCTCCAAAGACAGTTTTCCAAAACCTGAGGATCTCTGAGGTGCTTATGCTTACAGTG 1531
 Qy 435 CTGTCCTTATGGGATGTGTGCAGCTTTCTCAAGGCTCTCTGGGAGTGGGAGGCTGAAGA 494
 Db 1532 CTGTGCTTACGGCATCTGTGCGCAGCTTTCTCAAGACCTCTGCGGAGTGGCAGGCGGAGA 1591
 Qy 495 CTTTACCTTATGATGAGGAGTCTTCAAAAGGCGCTTGGGCTCTCTTGTGTCGACAGC 554
 Db 1592 CTTTATCCAGAGGAAGAGGAGGACCAAAAGGCGCTTGGTCTCTTGTGTCGACAGC 1651
 Qy 555 AGAGAACCACTATGACCAAGGAGCTGGATGAGCTCCAGCTGGAGATGGAGGAGCTCAAAGCC 614
 Db 1652 TGAGAACCACTATGACCTAGACCTGGATGAGCTCCAGATGGGACAGAGGACTCAAAGCC 1711
 Qy 615 ACACCCAGTGTCCAGTGTAGCCCTATCTCCAGGCGCTTCAA-GCCCTGTGAGTACCTCT 673
 Db 1712 ACACCCAGTGTCCAGTGTAGCCCTTGTTCAGGCGCTTCAAAGGCGCTTCCAGGACCTCT 1771
 Qy 674 TTGAAGCTGGGCGATCGGCTGCGCGTGGGCGCATGCTGCTCTCCGTCCTGCA 733
 Db 1772 TTGAGAGCTGGGCGATCGGCTTGTGTGGGCGCATGCTGCTCTCCGTCCTGTA 1831
 Qy 734 ATGACTGGTGTCTGTGACCTGTTCGCTGGGCGCTTCCGCTGCGGCGCTTCAAGT 793
 Db 1832 ACGGCTGGTGTCTGTGACCTGTTCGAGCGGAGCCAGCCGCTGTCCCGCTCAAGC 1891
 Qy 794 TTGTGTAGTGTGCAATTGACGCGCCAAACCTTGACTTGGCATTTCTGTGGCTTCTAG 853
 Db 1892 TTGTGTGGTGTGCGATGGCAGCGCCAAAGCGCTTGTGGGCGCATTTCTGTGTCTCTGG 1951
 Qy 854 CCTCAGTCGATGCGCTGACCTTTGCTCAGTTCTCTGAGTACGGAGCCGCTGGAGAGCG 913
 Db 1952 CCTCGGTGAGCGCTTGACCTATGCTGCTGAGTATGGAGCGCGCTGGGAGAGCG 2011
 Qy 914 GGCCTAGGCTGCGGCGGCACCTGGCTTCTTGGCAGTACTTGGGTTCGAGGACATCGGTGCTGC 973
 Db 2012 GTCTGGGCTGCCAGGCTACGGGCTTCTTGGCTGTCTTGGGTTCAGAGGCGTGGTGTGC 2071
 Qy 974 TGCTCACTCTGGCGCGAGTGCAGTCAGCGCTCTCGTCTCTCTGTCTCGGCGCTTATGGGA 1033
 Db 2072 TGCTCACACTGGCGCGGTGCAGTCAGTACATCTCGTGACCTGCGTCCGAGCTACGGGA 2131
 Qy 1034 AGTC-CCCTCTCTGGGCGAGGTTGCAGCAGGCGTCTAGCTGCTGCTGCTGCTGCTGCTG 1092
 Db 2132 AGGCGCGCTGCGCTGCGCAGCGTCCGCGCAGCGCACTTGGATGCTGCTGCTGCTGCTG 2191
 Qy 1093 CTGGCGCGCAGCTGCGCTGCGCTGCGGAGTACAGGCGCTTCCGCTGCTGCTGCTGCTGCTG 1152
 Db 2192 CTGGCGCGCAGCTGCGCTGCGCTGCGCTGCGGAGTATGGGCGCTTCCGCTGCTGCTGCTG 2251
 Qy 1153 CCCTACGCGCAGCTGAGGCTCAGCCAG-CHAGCCCTGGGCTTACCGTGGCGCTTGGTGTAT 1211
 Db 2252 CCCTACGCGCAGCTGAGGCTCAGCCAGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2311
 Qy 1212 GATGAACCTCTTCTGTTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1271
 Db 2312 GATGAACCTCTGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2371
 Qy 1272 GCGCGCGGCGAGCTTTCAGGCGCGTGGGAGCTGCGCATGCTGAGGAGCGTGGGCTGCTGCT 1331
 Db 2372 GCGAGGCGGAGCTTTCAGGCGCGTGGGAGCTGCGCATGCTGCGCAGCTGGGCTGCTGCTG 2431
 Qy 1332 CATCTTCGACAGCGGCTTCTTACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1391
 Db 2432 CATCTTCGAGATGGCGCTCTTACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2491
 Qy 1392 GGGCTCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1451
 Db 2492 GGGCTCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2551


```
REFERENCE
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: c9apbs-remail.nih.gov
Tissue Procurement: Lothar Hennighausen Ph.D., Priscilla Furth
Ph.D.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM10863 row: a column: 01
High quality sequence stop: 758.
Location/Qualifiers
1. .819
/organism="Mus musculus"
/mol_type="mRNA"
/strain="NMRI"
/db_xref="taxon:10090"
/clone="IMAGE:4934280"
/tissue_type="tumor, gross tissue"
/dev_stages="5 months"
/lab_host="DH10B"
/clone_lib="NCI CGAP Mam4"
/note="Organ: mammary; Vector: pCMV-SPORT6; Site 1: SalI;
Site 2: NotI; Cloned unidirectionally. Primer: Oligo dt.
Library constructed by Life Technologies. Investigators
providing samples: Lothar Hennighausen/Priscilla Furth,
NIH Reference for transgenic model: Li et al., Cell Growth
and Differentiation 7, 3-11 (1996)."
```

ORIGIN

```
Query Match 28.1%; Score 534.4; DB 4; Length 819;
Best Local Similarity 84.0%; Pred. No. 4.9e-110;
Matches 626; Conservative 0; Mismatches 117; Indels 2; Gaps 2;
```

QY 135 GGCCTCCACACACCGCATCTGGGAATTTGGAGCTNGACACCTTCAGCCAGCTGAGTCTC 194

DB 41 GGCCTCCACACATAACAGGATCAAGAAATTTGGTGC-AGATACCTTCAGCCAGCTGGGCTC 99

QY 195 CTTGCAAGCCCTGGATCTTAGCTGGAACGCCATCCGGTCCATCCACCTGAGGCCCTTCTC 254

DB 100 CTTGCAAGCTTAGACCTAGTTGGAATGCCATCCGTCGATCCACCTGAGGCTTTCTC 159

QY 255 CACCTTGCACTCCCTGGTCAAGCTGGACCTGACAGAACACCAAGCTGACACACTGCCCT 314

DB 160 AACCTTCGATCTTTGGTTAAGCTGGACCTGACTGACACCAACCACTGACACACTGCCCT 219

QY 315 GGCTGGACTTGGGGCTTGATGCATCTGAAGCTCAAGGGAACCTTGTCTCTCCAGGC 374

DB 220 GGCTGGCTGGAGGCTGATGCATCTGAAGCTCAAGGGAACCTTGGGCCCTGTCTCAGGC 279

QY 375 CTTCTCCAAAGACAGTTTCCAAAACCTGAGGATCTCTGGAGGTG-CCTTATGCTACCACT 433

DB 280 CTTCTCCAAAGACAGTTTCCAAAACCTGAGGATCTCTGGAGGTGCCCCCTACGCTACCACT 339

QY 434 GCTGTCCCTATGGGATGTGTGCCAGCTTTCTTCAAGGCCCTCTGGGAGTGGGAGGCTGAAG 493

DB 340 GCTGTGCTACGGCATCTGTGCCAGCTTCTTCAAGACCTCTGGGAGTGGGAGGCGAGG 399

QY 494 ACCTTACCTTGATGATGAGGAGTCTTCAAAAAGGCCCTTGGGCCCTCTTGGCAGACAAAG 553

DB 400 ACTTTTCATCCAGAAAGAGGAGGACCAAAAGAGGCCCTTGGGTCTCTTGTGTGACAAAG 459

QY 554 CAGAGAACCACTATACCAAGACCTGGATGAGCTCCAGCTGAGATGAGGAGCTCAAAAGC 613

DB 460 CTGAGAACCACTATACCTTAGACCTTGGATGAGCTCCAGATGGGACAGGAGCTCAAAAGC 519

QY 614 CACACCCAGTGTCTCCAGTGTAGCCCTACTCCAGGCCCTTCAAGGCCCTGTGAGTACCTCT 673

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

1 (bases 1 to 819)

NIH-MGC http://mgc.nci.nih.gov/.

National Institutes of Health, Mammalian Gene Collection (MGC)

Unpublished (1999)

Contact: Robert Strausberg, Ph.D.

Email: c9apbs-remail.nih.gov

Tissue Procurement: Lothar Hennighausen Ph.D., Priscilla Furth

Ph.D.

cDNA Library Preparation: Life Technologies, Inc.

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Incyte Genomics, Inc.

Clone distribution: MGC clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at:

http://image.llnl.gov

Plate: LLAM10863 row: a column: 01

High quality sequence stop: 758.

Location/Qualifiers

1. .819

/organism="Mus musculus"

/mol_type="mRNA"

/strain="NMRI"

/db_xref="taxon:10090"

/clone="IMAGE:4934280"

/tissue_type="tumor, gross tissue"

/dev_stages="5 months"

/lab_host="DH10B"

/clone_lib="NCI CGAP Mam4"

/note="Organ: mammary; Vector: pCMV-SPORT6; Site 1: SalI;

Site 2: NotI; Cloned unidirectionally. Primer: Oligo dt.

Library constructed by Life Technologies. Investigators

providing samples: Lothar Hennighausen/Priscilla Furth,

NIH Reference for transgenic model: Li et al., Cell Growth

and Differentiation 7, 3-11 (1996)."

ORIGIN

```
Query Match 28.1%; Score 534.4; DB 4; Length 819;
Best Local Similarity 84.0%; Pred. No. 4.9e-110;
Matches 626; Conservative 0; Mismatches 117; Indels 2; Gaps 2;
```

QY 135 GGCCTCCACACACCGCATCTGGGAATTTGGAGCTNGACACCTTCAGCCAGCTGAGTCTC 194

DB 41 GGCCTCCACACATAACAGGATCAAGAAATTTGGTGC-AGATACCTTCAGCCAGCTGGGCTC 99

QY 195 CTTGCAAGCCCTGGATCTTAGCTGGAACGCCATCCGGTCCATCCACCTGAGGCCCTTCTC 254

DB 100 CTTGCAAGCTTAGACCTAGTTGGAATGCCATCCGTCGATCCACCTGAGGCTTTCTC 159

QY 255 CACCTTGCACTCCCTGGTCAAGCTGGACCTGACAGAACACCAAGCTGACACACTGCCCT 314

DB 160 AACCTTCGATCTTTGGTTAAGCTGGACCTGACTGACACCAACCACTGACACACTGCCCT 219

QY 315 GGCTGGACTTGGGGCTTGATGCATCTGAAGCTCAAGGGAACCTTGTCTCTCCAGGC 374

DB 220 GGCTGGCTGGAGGCTGATGCATCTGAAGCTCAAGGGAACCTTGGGCCCTGTCTCAGGC 279

QY 375 CTTCTCCAAAGACAGTTTCCAAAACCTGAGGATCTCTGGAGGTG-CCTTATGCTACCACT 433

DB 280 CTTCTCCAAAGACAGTTTCCAAAACCTGAGGATCTCTGGAGGTGCCCCCTACGCTACCACT 339

QY 434 GCTGTCCCTATGGGATGTGTGCCAGCTTTCTTCAAGGCCCTCTGGGAGTGGGAGGCTGAAG 493

DB 340 GCTGTGCTACGGCATCTGTGCCAGCTTCTTCAAGACCTCTGGGAGTGGGAGGCGAGG 399

QY 494 ACCTTACCTTGATGATGAGGAGTCTTCAAAAAGGCCCTTGGGCCCTCTTGGCAGACAAAG 553

DB 400 ACTTTTCATCCAGAAAGAGGAGGACCAAAAGAGGCCCTTGGGTCTCTTGTGTGACAAAG 459

QY 554 CAGAGAACCACTATACCAAGACCTGGATGAGCTCCAGCTGAGATGAGGAGCTCAAAAGC 613

DB 460 CTGAGAACCACTATACCTTAGACCTTGGATGAGCTCCAGATGGGACAGGAGCTCAAAAGC 519

QY 614 CACACCCAGTGTCTCCAGTGTAGCCCTACTCCAGGCCCTTCAAGGCCCTGTGAGTACCTCT 673

Db 520 CAAACCCAGGTGTCAGTGCAGCCCTGTTCAGAGCCCTTCAAGCCCTGCAGACCTCT 579

QY 674 TTGAAAGCTGGGGCATCCGCTGCGCGTGTGGGCCATCGTGTGCTCTCCGCTCTGCA 733

Db 580 TTGAGAGCTGGGCATCCGCTTGTGTGTGGGCATCGTGTACTCTCCGCTACTCTGTA 639

QY 734 ATGGACTGGTCTGCTGACCGTGTTCGCTGGCGGCTGCCCCCTGCCCCCGGCTCAAGT 793

Db 640 ACGGCTGGTCTGCTGACAGTCTTTTGCAGCGGACCCAGCCGCTGCTCCCGTCAAGC 699

QY 794 TTGTGTAGTGGCATTCAGCGCGCAACACTTGACTGGCATTTCTCTGTGGCTTCTAG 853

Db 700 TTGTGTGGTGGCATTCAGCGCGCAACACTTGACTGGCATTTCTCTGTGGTCTCTGGG 759

QY 854 CCTCAGTCGATGCCCTGACCTTTTG 878

Db 760 CCTCGTGACCGCTTGACCTATGG 784

RESULT 6

LOCUS BG916782 812 bp mRNA linear EST 05-JUN-2001

DEFINITION 602816031P1 NCI_CGAP_Mam6 Mus musculus cDNA clone IMAGE:4935343 5',

ACCESSION BG916782

VERSION BG916782.1 GI:14297258

KEYWORDS EST.

SOURCE Mus musculus (house mouse)

ORGANISM Mus musculus

REFERENCE 1 (bases 1 to 812)

AUTHORS NIH-MGC http://mgc.nci.nih.gov/.

TITLE National Institutes of Health, Mammalian Gene Collection (MGC)

JOURNAL Unpublished (1999)

COMMENT Contact: Robert Strausberg, Ph.D.

Email: c9apbs-remail.nih.gov

Tissue Procurement: Jeffrey Green M.D.

cDNA Library Preparation: Life Technologies, Inc.

cDNA Sequencing by: Incyte Genomics, Inc.

Clone distribution: MGC clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at:

http://image.llnl.gov

Plate: LLAM10865 row: m column: 08

High quality sequence start: 5

High quality sequence stop: 792.

Location/Qualifiers

1. .812

/organism="Mus musculus"

/mol_type="mRNA"

/strain="FVB/N"

/db_xref="taxon:10090"

/clone="IMAGE:4935343"

/sex="female, virgin"

/tissue_type="infiltrating ductal carcinoma"

/dev_stages="5 months"

/lab_host="DH10B"

/clone_lib="NCI CGAP Mam6"

/note="Organ: mammary; Vector: pCMV-SPORT6; Site 1: SalI;

Site 2: NotI; Cloned unidirectionally. Primer: Oligo dt.

Library constructed by Life Technologies. Investigator

providing samples: Jeffrey Green, M.D., NIH"

ORIGIN

```
Query Match 27.2%; Score 516.2; DB 4; Length 812;
Best Local Similarity 81.3%; Pred. No. 6.4e-106;
Matches 635; Conservative 0; Mismatches 143; Indels 3; Gaps 3;
```

QY 391 TTCCCAAACTGAGGATCTCTGGAGTGCTTATGCTACCAAGTGTCTTCCCTATGGGATG 450

Db 1 TTCCCAAACTGAGGATCTCTGGAGTGCTTATGCTACCAAGTGTCTTCCCTATGGGATG 60

Db 1 TGTGACCTGCCACGGGGTGACTTTGAGGGCGGTGTGGAGCTGCGCCATGGTGGCCACCGTG 60

Qy 1324 GCCTGGCTCATCTTCGCAGACGGGCTCCTCTACTCTGCCGTGGCCCTCTCTCAGCTTCGCC 1383

Db 61 GCCTGGCTCATCTTCGCAGAGGCTCCTCTACTCTGCCGTGGCCCTCTCTCAGCTTCGCC 120

Qy 1384 TCCATGCTGGGCTCTTCCCTGTGTCAGCCCGAGGCGCTCAAGTCTGCTCTGCTGTGGTG 1443

Db 121 TCTATGCTGGGCTCTTCCCTGTGTCACCCCGAGGCTCAAGTCAAGTCTCTGCTGTGGTG 180

Qy 1444 CTGCCCCCTGCTGCTGCTCAACCACTGCTGTACCTGCTCTTCAACCCCACTTCGG 1503

Db 181 CTGCTCTGCTGCTGCTCAACCACTGCTGTACCTGCTCTTCAACCCCACTTCGG 240

Qy 1504 GATGACCTTCGGCGCTTCGGCGCGCGAGGAGCTCAGGGCCCTAGCCTATCTGCG 1563

Db 241 GATGACCTTCGGCGCTTCGGCGAGCCTCGGTCCCGAGGCGCTTAGCCTAGCTGCA 300

Qy 1564 GCGGGGAGCTGGAGAAGCTCTCTGTGATCTACCCAGGCGCTGTAGCTTCTCTGAT 1623

Db 301 GCGGCTGAGCTGGAGAAGCTCTCTGCACTCCACCAAGCGCTGTGCTCTCTCAGAT 360

Qy 1624 GTGATCTCATCTCTGGAAGCTCTCAAGCTGGCGGCGCCCTGGCTGAGACCTATGCG 1683

Db 361 GTGATCTCATCTCTGGAAGCTCTGAGGCTGGCGAGCCTCTGCTGAGACCTATGCG 420

Qy 1684 TTCCCTCAGTGACCTCATCTCTCTGTCAGAGCGAGGCGCCCGAGGCTGGAGGCGAGC 1743

Db 421 TTCCCTCAGTGACCTCATCTCTCTGTCAGACATCAGCGGGGCTTACAGCTGGAGGGAAC 480

Qy 1744 CATTGTGTAGAGCCAGAGGGAACCACTTTGGGAACCCCAACCTCTCATGATGAGAA 1803

Db 481 CATTGTGTAGAGCTGTGATGGAACCAAGTTTGGGAACCCCAACCTCTCATGAGGAGAA 540

Qy 1804 CTGCTGCTGAGGCGAGGAGTCTACGCGAGGCTGAGGCTTCTCAGGGGTGGCGC 1863

Db 541 CTGCTGCTGAGGCGAGGAGGAGCCACTTTGGCAGGCTGTGGCTCTTCGTGGGTGGAGCC 600

Qy 1864 TTTCAGCCCTCTGGCTTGGCTTTGCT 1890

Db 601 CTCTGGCCCTCTGGCTCTCTCTTGCT 627

RESULT 8

AK040883

LOCUS

DEFINITION Mus musculus adult male aorta and vein cDNA, RIKEN full-length enriched library, clone:A530037C04 product:CDNA FLJ14471 FIS, CLONE MAMMAL001030, WEAKLY SIMILAR TO LUTROPIN-CHORIOGONADOTROPIC HORMONE RECEPTOR homolog [Homo sapiens], full insert sequence.

ACCESSION AK040883

VERSION AK040883.1 GI:26088131

KEYWORDS HTC; CAP trapper.

SOURCE Mus musculus (house mouse)

ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE 1 Carninci,P. and Hayashizaki,Y. High-efficiency full-length cDNA cloning Meth. Enzymol. 303, 19-44 (1999)

JOURNAL 99279253

MEDLINE 10349636

PUBMED 10349636

REFERENCE 2 Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K., Itoh,M., Konno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y. Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes Genome Res. 10 (10), 1617-1630 (2000)

JOURNAL 20499374

MEDLINE 11042159

PUBMED 11042159

REFERENCE 3 Shibata,K., Itoh,M., Aizawa,K., Nagaoka,S., Sasaki,N., Carninci,P.,

Konno,H., Akiyama,J., Nishi,K., Kiteunai,T., Tashiro,H., Itoh,M., Sumi,N., Ishii,Y., Nakamura,S., Hazama,M., Nishine,T., Harada,A., Yanamoto,R., Matsumoto,H., Sakaguchi,S., Ikegami,T., Kashiwagi,K., Fujiwaka,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E., Watahiki,M., Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Matsuura,S., Kawai,J., Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A. and Hayashizaki,Y. RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer Genome Res. 10 (11), 1757-1771 (2000)

20530913

11076861

4

The RIKEN Genome Exploration Research Group Phase II Team and the FANTOM Consortium.

FUNCTIONAL annotation of a full-length mouse cDNA collection Nature 409, 685-690 (2001)

5

The FANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team.

Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs Nature 420, 563-573 (2002)

6 (bases 1 to 3855)

Adachi,J., Aizawa,K., Akimura,T., Arakawa,T., Bono,H., Carninci,P., Fukuda,S., Furuno,M., Hanagaki,T., Hara,A., Hashizume,W., Hayashida,K., Hayatsu,N., Hiramoto,K., Hiraoka,T., Hirozane,T., Hori,F., Imotani,K., Ishii,Y., Itoh,M., Kagawa,I., Kasukawa,T., Katoh,H., Kawai,J., Kojima,Y., Kondo,S., Konno,H., Kouda,M., Koya,S., Kurihara,C., Matsuyama,T., Miyazaki,A., Murata,M., Nakamura,M., Nishi,K., Nomura,K., Numazaki,R., Ohno,M., Ohsato,N., Okazaki,Y., Saito,R., Saitoh,H., Sakai,C., Sakai,K., Sakazume,N., Sano,H., Sasaki,D., Shibata,K., Shinagawa,A., Shiraki,T., Sogabe,Y., Tagami,M., Tagawa,A., Takahashi,F., Takaku-Akahira,S., Takeda,Y., Tanaka,T., Tomaru,A., Toya,T., Yasunishi,A., Muramatsu,M. and Hayashizaki,Y.

Direct Submission

Submitted (16-JUL-2001) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail:genome-res@gsc.riken.jp, URL:http://genome.gsc.riken.jp/, Tel:81-45-503-9222, Fax:81-45-503-9216)

CDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in Riken. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.

Please visit our web site for further details.

URL:http://genome.gsc.riken.jp/

URL:http://fantom.gsc.riken.jp/.

Location/Qualifiers

1..3855

/organism="Mus musculus"

/mol_type="mRNA"

/strain="C57BL/6J"

/db_xref="FANTOM,DB:A530037C04"

/db_xref="taxon:10090"

/clone="A530037C04"

/sex="male"

/tissue_type="aorta and vein"

/clone_lib="RIKEN full-length enriched mouse cDNA library"

/dev_stage="adult"

<1..642

misc_feature

note="CDNA FLJ14471 FIS, CLONE MAMMAL001030, WEAKLY SIMILAR TO LUTROPIN-CHORIOGONADOTROPIC HORMONE RECEPTOR homolog [Homo sapiens] (SPTR|096K69, evidence: FASTY, 85.8%ID, 92.1%length, match=637) putative"

ORIGIN

Query Match 25.3%; Score 479.8; DB 3; Length 3855;

Best Local Similarity 85.4%; Pred. NO. 1.4e-97;

cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov> row: f column: 13
Plate: LLMW10865 row: f column: 13
High quality sequence stop: 644.
Location/Qualifiers

FEATURES

source

1..753
/organism="Mus musculus"
/mol_type="mRNA"
/strain="FVB/N"
/db_xref="taxon:10090"
/clone="IMAGE:4935180"
/sex="female, virgin"
/tissue_type="infiltrating ductal carcinoma"
/dev_stage="5 months"
/lab_host="DH10B"
/clone_lib="NCI CGAP Mam6"
/note="Organ: mammary; Vector: pCMV-SPORT6; Site 1: SalI;
Site 2: NotI; Cloned unidirectionally. Primer: Oligo dt.
Library constructed by Life Technologies. Investigator
providing samples: Jeffrey Green, M.D., NIH"

ORIGIN

Query Match 24.1%; Score 457.6; DB 4; Length 753;
Best Local Similarity 83.8%; Pred. No. 1.2e-92;
Matches 612; Conservative 0; Mismatches 109; Indels 9; Gaps 8;
QY 832 GGCATTCTGTGGCTTCTAGCTCAGTCGATGCCCTGACCTTTGGTCAGTTCTCTGAG 891
Db 28 GGCATTCTGTGGTCTCTGGCTCGGTGACGCGCTTGACCTATGTCAGTTCGCTGAG 87
QY 892 TACGAGCCCTGGGAGACGGGTAGCTGCGCGGCCACTGGCTTCTGCGCACTTT 951
Db 88 TATGAGCCCGTGGGAGAGCGGTCTGGCTGCCAGGCTACGGGCTTCTGCTGCTCTG 147
QY 952 GGTGCGAGGATCGGTCTGCTGCTCACTCTGGCGGAGTGCAGTGCA-GCGTCTCCGT 1010
Db 148 GGTTCAGAGGCGTGGTGTCTGTCTACACTGGCGGGTGCATGTCATGCTCTGT 207
QY 1011 CTCCTGTGTGGGGCTATGGAAGTCCCTCTCTGCGGAGCTTCGAGCAGGGGTCT 1070
Db 208 GACCTGCTCGAGCTACTCGGAGAGCGCGTCTGCGCTGCGCAGCGTCCGCGAGGCGCACT 267
QY 1071 AGGCTGCTGGC--ACTGGCAGGCTGGCCGCGC-CGACTGCCCCCTGGCTCTAGTGGGAGA 1127
Db 268 GGGATGCTTGGCTGTGTGGCCGTGGCTGGCCGAAGCACTGGCGCTGGGCTCGGTGGGAGA 327
QY 1128 ATACGGGCGCTCCCGCACTCTGCTGCCCTACGCGCCAC-CTGAGGGTCAGCCAGCGGCC 1186
Db 328 GTATGGCGCTCCCGCACTCTGCTGCCCTACGCGCCACACCGAGGGCGCTGGGACGCC 387
QY 1187 TGGGCTTACCGTGGCCCTGTGATGATGAATCTCTTCTGTTTCTGTCGTCGTCGCGGTG 1246
Db 388 TGGGCTTCTGCTGTAGCCCTGGTGTGATGAATCTGCTCTGCTCTCTGCTGGTGGCGCGC 447
QY 1247 CCTACATCAAACTGTACTGTGACCTCGCGCGGGCGACTTTGAGCCCTGTGGGACTCGG 1306
Db 448 CCTACATCAAGCTCTACTGTGACCTGTCACTGGGGTGACTTTGAGCCCGTGGGACTCGG 507
QY 1307 CCATGGTGAGGACGTGGCTGGCTCATCTTCGACAGAGGGGCTCTCTACTGTCTCCGTGG 1366
Db 508 CCATGGTGCGCACGTGG-CTGGCTCATCTTTGCGAGATGGCTCTCTACTGCCCCGTGG 566
QY 1367 CTTCTCTCAGCTTCGCTTCCATGCTGGGCTCTTCCCTCTGTACGCCCCGAGGCCGCTCAAGT 1426
Db 567 CTTCTCTCAGCTTTTGGCTCTATGCTGGGCTCTTCCCTGTCTAC-CCCGAGGCTGTCAAGT 625
QY 1427 CTGCTCTGCTGGTGTGCTGCCCCCTGCTGCTGCTCAACCCACTGCTACTGCTGCTCT 1486
Db 626 CAGTCTTCTGCTGGTGTGCTGCTGCTGCTGCTGCTCAACCCACTGCTCTA-CTGCTCT 684

QY 1487 TCAACCCCACTTCGGGATGACCTTCGGCGCTTCGGCCCGCGCAGGGAATCAGGCG 1546
Db 685 TCAACCTCTATTCGGGGATGA-CTTCGGCGCTCTGCGCAAGCCTCGGAACCAAGGC 743
QY 1547 CCCTAGCCTA 1556
Db 744 CCCTAGCCTA 753

RESULT 11

BF158974
LOCUS BF158974
DEFINITION 601766476F1 NCI_CGAP_Lu29 Mus musculus cDNA clone IMAGE:3982506 5',
mRNA sequence.
ACCESSION BF158974
VERSION BF158974.1 GI:11039068
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 869)
AUTHORS NIH-MGC <http://mgs.nci.nih.gov/>
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: Gilbert Smith, Ph.D.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
Plate: LLMW181 row: g column: 19
High quality sequence stop: 640.
Location/Qualifiers

FEATURES

source

1..869
/organism="Mus musculus"
/mol_type="mRNA"
/strain="CZECH II"
/db_xref="taxon:10090"
/clone="IMAGE:3982506"
/tissue_type="spontaneous tumor, metastatic to mammary."
Stem cell origin.
/lab_host="DH10B"
/clone_lib="NCI_CGAP_Lu29"
/note="Organ: lung; Vector: pCMV-SPORT6; Site 1: SalI;
Site 2: NotI; Cloned unidirectionally. Primer: Oligo dt.
Library constructed by Life Technologies. Investigator
providing samples: Gilbert Smith, NIH"

ORIGIN

Query Match 22.3%; Score 424.4; DB 2; Length 869;
Best Local Similarity 81.9%; Pred. No. 3.9e-85;
Matches 549; Conservative 0; Mismatches 116; Indels 5; Gaps 5;
QY 522 AAAAGGCCCCCTGGCCCTCTTTCGACAGCAGAGAACCACTATGACGAGCCTGA 581
Db 1 AAAGAGGCCCCCTGGCTCTCTTGTGGACAAGCTGAGAACCACTATGACTA-GACCTGA 59
QY 582 TGAGCTCCAGCTGGAGATGGAGGACTCAAGCCACACCCAGTGTCCAGTGTAGCCCTAC 641
Db 60 TGAGCTCCAGATGGGACAGAGGACTCAAGCCAAACCCAGTGTCCAGTGCAGCCCTGT 119
QY 642 TCCAGGCCCCCTTCAAGCCCTGTGAGTACCTCTTTTAAAGCTGGGGCAATCCGCCCTGGCCGT 701
Db 120 TCCAGGCCCCCTTCAAGCCCTGCGAGCAGCTCTTTGAGAGCTGGGCACTCCGCCCTTGTGT 179
QY 702 GTGGGCCAATCGTGTGTCTCCGTGCTCTGCAATGGAGTGTGCTGTGACCGGTGTTGCG 761
Db 180 GTGGGCCAATCGTGTGTCTCCGTGCTCTCCGTACTGTGTAAAGGGCTGTGTGTGACAGTCTTTTGC 239

Qy 762 TGGGGGCTGCTCCCTGCCCCCTGAGTTTGTAGTGCAGTTGCAGGCGCAA 821
Dy 240 CAGCGGACCCAGCGCTGTCCCGTCAAGCTTGTGTGGTGCAGTGGCAGGCGCA 299
Qy 822 CACCTTGACTGGCATTTCTGTGGCTTCTAGCCTCAGTCGATGCCCTTGAGCTTTGTGCA 881
Dy 300 CGCCCTGACGGGCAATTTCTGTGGTCTCTCGCTCGGTGGACGCTTGACCTATGGTCA 359
Qy 882 GTTCTCTAGTACGGAGCCCTGGGAGACGGGCTAGGCTGCCGGGCACCTGCTTCCT 941
Dy 360 GTTCTCTAGTACGGAGCCCTGGGAGAGGGCTTGGGCTGCCAGGCTACGGGCTTCCT 419
Qy 942 GGCAGTACTTGGGTGGAGGAGCATCGTCTGCTCTCACTTGGCGCGAGTGCAGTGCGAG 1001
Dy 420 GGCTGTCTGGTTCAGAGGGCTGGTGTCTGCTCTCACTTGGCGCGGTCAGTGCGAG 479
Qy 1002 CGTCTCGCTTCTGTGTGCGGGCTATGGGAAGTCCCTTCCCTGGGAGAGCTTGCAGC 1061
Dy 480 CATCTCTGTGACCTG-GTCCGAGCCTACGGGAAGGGCGCTGCTGG-CAGCGTCCGCGC 537
Qy 1062 AGGGGTCTAGGCTGCTGGACATGGCAGGCTGGCGCCGACCTGCCCCCTGGCTCAGT 1121
Dy 538 AGGGGCACTGTGATGCTGCTGGCGCTGGCGGCTGGCCTGCGACACTG-CGCTGGGCTCAGT 596
Qy 1122 GGGAGATACGGGGCTCCCGCACTGCTGCTGCTTCCCTACGCGCA-CCTGAGGGTCAGCCAG 1180
Dy 597 GGGAGATATGGCGCTCCCGCACTGCTGCTGCTTCCCTAAGCCCCACCGGAGTGGCGGG 656
Qy 1181 CAGCCCTGGG 1190
Dy 657 CGGCCCTGGG 666

RESULT 12
BE149368/c 461 bp mRNA linear EST 21-JUN-2000
LOCUS RCI-HT0255-151099-011-d07 HT0255 Homo sapiens cDNA, mRNA sequence.

DEFINITION BE149368
ACCESSION BE149368
VERSION BE149368.1 GI:8612089
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 461)
Dias Neto, E., Garcia Correa, R., Verjovski-Almeida, S., Briones, M.R.,
Nagai, M.A., da Silva, W. Jr., Zago, M.A., Bordin, S., Costa, F.,
Goldman, G.H., Carvalho, A.P., Matekuma, A., Baia, G.S., Simpson, D.H.,
Brunstein, A., deOliveira, P.S., Bucher, P., Jongeneel, C.V.,
O'Hare, M.J., Soares, F., Brentani, R.R., Reis, L.F., de Souza, S.J. and
Simpson, A.J.

TITLE Shotgun sequencing of the human transcriptome with ORF expressed
sequence tags
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
MEDLINE 20202663
PUBMED 10737800
COMMENT Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome
Project. This entry can be seen in the following URL
(http://www.ludwig.org.br/scripts/gethtml2.pl?tl=et2-RC1-HT0255-151
039-011-d07&t3=1999-10-15&t4=1)
Seq primer: puc 18 forward
High quality sequence start: 23
High quality sequence stop: 461.
Location/Qualifiers

FEATURES

source 1. .461
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/dev_stage="Adult"
/clone_lib="HT0255"
/note="Organ: head neck; Vector: puc18; Site 1: SmaI;
Site 2: SmaI; A mini-library was made by cloning products
derived from ORESTES PCR (U.S. Letters Patent application
No. 136,716 - Ludwig Institute for Cancer Research)
profiles into the puc 18 vector. Reverse transcription of
tissue mRNA and cDNA amplification were performed under
low stringency conditions."

ORIGIN
Query Match 20.3%; Score 385.2; DB 2; Length 461;
Best Local Similarity 97.2%; Pred. No. 2.6e-76;
Matches 445; Conservative 0; Mismatches 8; Indels 5; Gaps 5;
Qy 801 AGGTGGCATTTGCGGGCCCAACACCTTGACTGGCATTTCTGTGGCCTTCTAGCCTCAGT 860
Dy 457 AGGCTTGAATTCGAGGCGCCCTACACCTTGACTGGCATTTCTGTGGCCTTCTAGCCTCAGT 398
Qy 861 CGATGCCCTCAGCTTTTGGTTCAGTTCTCTGAGTACGGAGCCCGCTGGGAGACGGGGCTAGG 920
Dy 397 CGATGCCCTCAGCTTTTGGTTCAGTTCTCTGAGTACGGAGCCCGC-GGGAGACGGGGCTAGG 339
Qy 921 CTGCGGGGCCACCTGGCTTCTGAGTACCTTGGGTGGAGGATCGGTGCTGTGCTCAGT 980
Dy 338 CTGCGGGGCCACCTGGCTTCTGAGTACCTTGGGTGGAGGATCGGTGCTGTGCTCAGT 279
Qy 981 TCTGGCGCAGTGCAGTGCAGCTTCCGCTCTCTGTGTCGGGGCTATGGGAGTCCCC 1040
Dy 278 TCTGGCGCAGTGCAGTGCAGCTTCCGCTCTCTGTGTCGGGGCTATGGGAGTCCCC 219
Qy 1041 CTCCTGGGCGAGGTTTCGAGCAGGGGCTCTAGGTCCTGGCTGCTGGCCTATGGCAGGCTGGCGCG 1100
Dy 218 CTCCTGGGCGAGGTTTCGAGCAGGGGCTCTAGGTCCTGGCTGCTGGCCTATGGCAGGCTGGCGCG 159
Qy 1101 CGCACTGCCCTGCGCTCAGTGGAGAAATACGGGGCTCCCGCCTCTGCTGCTGCTGCTGCTGCTGCT 1160
Dy 158 CGCACTG-CCTTGGCTCAGTGGGAGAAATACGGGGCTCCCGCCTCTGCTGCTGCTGCTGCTGCTGCT 100
Qy 1161 GCCACCTGAGGGTCAGCCAGCAGCCCTGGGCTTTCACGTCGGCGCTGCTGCTGCTGCTGCTGCTGCT 1220
Dy 99 GCCACCTGAGGGTCAGCCAGCAG-CCTGGGCTTCCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 42
Qy 1221 CTCTCTGTTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1258
Dy 41 CTCTCTGTTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 5

RESULT 13
AY411734 2724 bp DNA linear GSS 16-DEC-2003
LOCUS AY411734
DEFINITION Mus musculus GPR49 gene, VIRTUAL TRANSCRIPT, partial sequence,
genomic survey sequence.
ACCESSION AY411734
VERSION AY411734.1 GI:39767702
KEYWORDS GSS.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 2724)
REFERENCE
AUTHORS Clark, A.G., Glanowski, S., Nielson, R., Thomas, P., Kejariwal, A.,
Todd, M.A., Tanenbaum, D.M., Civello, D.R., Lu, F., Murphy, B.,
Fieriera, S., Wang, G., Zheng, X.H., White, T.J., Shinsky, J.J.,
Adams, M.D. and Cargill, M.
TITLE Inferring nonneutral evolution from human-chimp-mouse orthologous
gene trios
JOURNAL Science 302 (5652), 1960-1963 (2003)
PUBMED 14671302

REFERENCE	2 (bases 1 to 2724)
AUTHORS	Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A., Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B., Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J., Adams,M.D. and Cargill,M.
TITLE	Direct Submission
JOURNAL	Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive, Rockville, MD 20850, USA
COMMENT	This sequence was made by sequencing genomic exons and ordering them based on alignment.
FEATURES	Location/Qualifiers
source	1..2724 /organism="Mus musculus" /mol_type="genomic DNA" /db_xref="taxon:10090" <1..>2724 /genes="GPR49" /locus_tag="HCM4322"
gene	
ORIGIN	
Query Match	19.1%; Score 362.4; DB 9; Length 2724;
Best Local Similarity	56.4%; Pred. No. 4.9e-71;
Matches	769; Conservative 0; Mismatches 567; Indels 27; Gaps 4;
Qy	161 AATTGGAGCTNGACACCTTCAGGCAGCTGAGCTCCCTGCAAGCCCTCGATCTCTAGCTGGA 220
Db	1163 AAATTAAGGCGAGCACCTTTTCAGCAGTTGTTTAACTCCGATCTCTGAACTTAGCATGGA 1222
Qy	221 ACGCCATCCGGTCCATCCACCTTGAGCCCTTCTCCACCCTGCACCTCCCTCGTCAAGCTGG 280
Db	1223 ATAAATTTGCTATCATTCACCCCAATGCGTTTCTACGTTGCCGTCTCTAATAAAGTTGG 1282
Qy	281 ACCTGACAGACAACACAGCTGACCACACTGCCCTCGCTGGCTGGACTTGGGGGCTTGATGCATC 340
Db	1283 ACCTATCATCCATCTCTGCTGCTCTCCCTGTGACTGGTTACATGGTTTAATCTCACT 1342
Qy	341 TGAAGCTCAAGGGAACCTTGCTCTCTCCAGGCGCTTCTCCAGGACAGTTCCTCCCAAAAC 400
Db	1343 TAAAAATTAAACAGGGAACCGAGCCTTACAGAGCCTTGATACCATCTGCAAACTTCCAGAGC 1402
Qy	401 TGAGGATCCTGGAGGTGCCCTTATGCCTACAGTGCCTGCTCCCTATGGGATGTGTGCCAGCT 460
Db	1403 TCAAGATTATAGAAATGCCATCTGCTTACCAGTGTGTGTGATTTGGGGGGTGTGAGAATG 1462
Qy	461 TCTTCAAGGCGCTCTGGCAGCTGGGAGGCTGAAAGACCTTTCACCTTGATGTAGGAGAGTCTT 520
Db	1463 TCTATAAATTTCTAACCAATGNAATAAGACGACGACAGTGTGACGACCTTCATA 1522
Qy	521 CAAAAAGGCCCTCGGGCCTCCTTGCCAGACAAGCAGAGAACCACTATGACACGAGACCTGG 580
Db	1523 AGAAAGACGCTGGGTATTATTTCAAGT-----TCAAGATGAGCGGACCTTG 1567
Qy	581 ATGAGCTCCAGCTGGAGAT---GGAGGACTCAAGCCACACCCCACTGTCAGCTGATGACC 637
Db	1568 AAGATTTCCTACTTGACTTTTGAGGAAGACCTGAAGGCCCTTCACTCGGTGCAGTGTCTCGC 1627
Qy	638 CTACTCCAGGCCCTTCAAGCCCTGTGAGTACCTCTTTGAAAAGCTGGGGGCATTCGCGCTGG 697
Db	1628 CTTCCCGAGTCCCTTCAAGCCCTGTGAGCACCTATTTTGGTAGCTGCTGATCCGAATCG 1687
Qy	698 CCGTGTGGGCCATCGTGTTCCTTCGCTGTGCAATGGACGTGGTCTGCTGACCGGTG 757
Db	1688 GGGGTGTGGACCAAGGACGAGTACTGCGCGCTTTCCTGCAATGCTTGGTGGCTTTGACCGGTG 1747
Qy	758 TCGTGGCGGGCCCTGCCCCCGTCAAGTTTGTGTGATAGTGCATTTGCAGGTGCAGGCG 817
Db	1748 TC---AGAACTCCTCTGTACATCTCTTCCATAAAGCTGCTAAATGGGGGTAAATCGCGGTAG 1804
Qy	818 CCAACACCTTGTACTTGGCATTTCTGTGGCGCTTCTAGCCTCAGTTCGATGCCTGACCTTTG 877
Db	1805 TGGACATTCATAGGGGGTCTCAGTCTGTGCTGGCTGCGGTGGATGCAATTCATCTTTG 1864
Qy	878 GTCAGTTCTCTGAGTACGGAGCCCGCTGGGAGACGGGGCTTAGGCTGCCGGGCCCACTGGCT 937

COMMENT This sequence was made by sequencing genomic exons and ordering them based on alignment.

FEATURES source Location/Qualifiers

1..2724

/organism="Homo sapiens"

/mol_type="genomic DNA"

/db_xref="taxon:9606"

<1..>2724

/gene="GPR49"

/locus_tag="HCW4322"

ORIGIN

Query Match 18.7%; Score 355.6; DB 9; Length 2724;

Best Local Similarity 56.2%; Pred. No. 1.7e-69;

Matches 781; Conservative 0; Mismatches 580; Indels 28; Gaps 5;

Qy 135 GGCTCCAAACACACCGCATCTGGAAATTTGGAGCTNGACACCTTCAGCCAGCTGAGCTC 194

Db 1138 GACCTAAGACATAAATGAAATCTAGCAATTTAAAG-TTGACACTTTCCAGCAGTTGCTTAG 1196

Qy 195 CCTGCAAGCCCTGGATCTTAGCTGGAAGCCATCGGTCCATCCACCTGAGGCTTCTC 254

Db 1197 CCTCGATCGCTGAATTTGGCTTGGAAACAAATTTGCTAATTAATCACCCCAATGCATTTTC 1256

Qy 255 CACCTCGACATCCCTGGTCAAGCTGGACCTGACAGACAAACAGCTGACCACTGCCCT 314

Db 1257 CACTTTGCCATCCCTAATAAGCTGGACCTATGTCCAACCTCTCTGTGCTTTTTCCTAT 1316

Qy 315 GGCTGGACTTGGGGCTTGTATGCACTCTGAAGCTCAAAAGGGAACCTTGTCTCTCCAGGC 374

Db 1317 AACTGGGTACATGGTTTAACCTACTTAAATTAACAGGAAATCATGCTTACAGAGCTT 1376

Qy 375 CTTCTCCAAAGACAGTTTCCAAACTGAGATCTCGGAGTGTCTTATGCTTACCACTG 434

Db 1377 GATATCATCTGANAACCTTCCAGAACTCAAGGTTATAGAAATGCTTTATGCTTACCACTG 1436

Qy 435 CTGTCCCTATGGGATGTGTGCAGCTTCTTCAAGCCCTCTGGCAGTGGGAGGCTGAAGA 494

Db 1437 CTGTGCAATTGGAGTGTGTAGATGCTATAGATTTCTAATCAATGGAATAAAGTGA 1496

Qy 495 CCTTCACCTTGATGATGAGGAGTCTTCAAAAGGCCCTTGGGCCCTCTTGCACAGCAAGC 554

Db 1497 CAACAGCAGTATGACGACCTTCTATAAGAA-----GATGCTGAATGTT 1541

Qy 555 AGAGAACCACTATACACGAGCCTGGATGAGCTCCAGCTGGAGAT---GGAGGACTCAA 611

Db 1542 TCAGGCTCAAGATGAACGTGACCTTGAAGATTTCTCTGCTGACTTTGAGGAAGACCTGAA 1601

Qy 612 GCCACACCCAGTGTCCAGTGTAGCCCTACTCCAGGCCCTTCAAGGCCCTGTGAGTACCT 671

Db 1602 AGCCCTTCATTCAGTGAGTGTTCACCTTCCCGAGGCCCTTCAACCCCTGTGACACCT 1661

Qy 672 CTTTGAAGCTGGGGCATCCGCCCTGGCCGTGTGGCCATCGTGTGTCTCTCCGTGCTG 731

Db 1662 GCTTGATGGCTGGCTGATCAGAAATGGAGTGTGACCATAGACAGTCTTGGCAGCTTACTTG 1721

Qy 732 CAATGGAGTGTGTGCTGCTACCGTGTTCGCTGGGGGCTGCCCCCTGCCCCCGGTCAA 791

Db 1722 TAATGCTTTGGTGTACTTCAACAGTGTTCAGATCC---CCTGTGTACATTTCCCCCATTA 1778

Qy 792 GTTTGTGTAGTGTCCGATTTGAGCGGCCAACACCTTGACTGGCATTTCTGTGGCTTCT 851

Db 1779 ACTGTTAATTTGGGTCTATCGCAGCAGTGAACATGCTCAGGGAGTCTCCAGTGGCTGCT 1838

Qy 852 AGCTTCAGTCGATCCCTGACCTTTTGTGTGAGTCTCTGAGTACGAGCCCGCTGGAGAC 911

Db 1839 GGCTGGTGTGATGCTGCTTCACTTTTGGCAGCTTTGACGACATGCTGCTGGTGGAGAA 1898

Qy 912 GGGCTAGGCTGCGCGGCCACTGCTTCTCGGAGTACTTGGGTGCGAGGATCGGTGCT 971

Db 1899 TGGGGTGGGTGGCCATGCTATGTTTTTTTGTCCATTTTGTCTTTCAGAAATCATCTGTTT 1958

Qy 972 GCTGCTCACTCTGGCCGACGTGCACTGACGGTCTCCGTCTCTCTGTGTCGGGGCTATGG 1031

Db 1959 CCTCTTACTCTGSCAGCCCTGGAGCGTGGGTCTCTGTGAAATATTCTGCAAAATTTGA 2018

Qy 1032 GAAGTCCCCCTCCCTGGGAGCGTTCGAGCAGGGGTCTTAGCTGCGCTGGCACTGGCAGG 1091

Db 2019 AACGAAAGCTCCATTTTCTAGCCTGAAAGTAATCATTTTGTCTGTGCTGCTGCTGCT 2078

Qy 1092 GCTGGCGCGCAGCTGCCCTGGCCTCAGTGGGAGAAATACGGGGCCTCCCACTCTGCT 1151

Db 2079 GACCATGGCGCAGTTCCTCTGCTGGGTGGCAGCAAGTATGGCGCCTCCCTCTCTGCT 2138

Qy 1152 GCCCTAGCGCCCACTCGAGGCTCAGCCAGCAGCCCTGGGCTTCCACCTGGGCCCTGGTGAT 1211

Db 2139 GCCTT-----TGCTTTTGGGAGCCAGCACCATGGGCTACATGCTCTCATCTT 2192

Qy 1212 GATGAATCTCTTCTGTTTCTGTTGCTGGCGGTGCTTACATAACATGTAATCTGACCT 1271

Db 2193 GCTCAATTCCTTTGCTTCTCATGATGACCATTTGCCCTACACCAAGCTCTACTGCAATTT 2252

Qy 1272 GCCCGGGGCGACTTTGAGCCGCTGTGGGACTGGCCATGCTGAGGACGCTGGCCTGGCT 1331

Db 2253 GGACAAGGGAGACCTGGAGAAATTTGGGACTGCTCTATGGTAAACACATTTGCCCTGTT 2312

Qy 1332 CATCTTCGACAGCGGCTCCTCTACTGTCCCGTGGCCTTCTCAGCTTTCGCTCCATGCT 1391

Db 2313 GCTCTTCAACCACTGCACTCTAACTGCCCTGTGGCTTCTTGTCTCTCTCTCTTAAT 2372

Qy 1392 GGGCCTCTTCCCTGTACGCCCGAGCGCTCAAGTCTGTCTGTGTGTGTGTGCTGCT 1451

Db 2373 AAACCTTACATTTATCAGTCTGAAATTAAGTTTATCTCTGTGGTGTGTGCTGCT 2432

Qy 1452 GCCTGCTGCTCAACCCACTGCTGTACCTGCTTCAACCCCACTTCCGGGATGACCT 1511

Db 2433 TCCTGCAATGCTCAATCCCTTCTCTACATCTTCTCAATCCTCACCTTTAAGGAGATCT 2492

Qy 1512 TCGGCGGCT 1520

Db 2493 GGTGAGCT 2501

RESULT 15

CO351740

LOCUS DR_AOV_NRM05_F12 adult ovary normalized (TILL) Danio rerio cDNA, 640 bp mRNA linear EST 29-JUN-2004

DEFINITION mRNA sequence.

ACCESSION CO351740

VERSION CO351740.1 GI:49433057

KEYWORDS EST.

SOURCE Danio rerio (zebrafish)

ORGANISM Danio rerio

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes; Cyprinidae; Danio.

REFERENCE 1 (bases 1 to 640)

AUTHORS Li, Y., Chia, J.M., Bartfai, R., Christoffels, A., Yue, G.H., Ke, D., Ho, M.Y., Hill, J.A., Stupka, E. and Orban, L.

TITLE Comparative analysis of the testis and ovary transcriptomes in zebrafish by combining experimental and computational tools

JOURNAL Unpublished (2004)

COMMENT Contact: Laszlo ORBAN

Reproductive Genomics Group

Tamashek Life Sciences Laboratory

1 Research Link, the NUS, Singapore 117 604

Tel: +65 6872 7413

Fax: +65 6872 7007

Email: laszlo@tll.org.sg

Similar to IGR6 HUMAN

High quality sequence stop: 640.

Location/Qualifiers

1..640

/organism="Danio rerio"

/mol_type="mRNA"

/strain="AB"

```

/db_xref="taxon:7955"
/sex="female"
/dev stage="adult (fully mature)"
/clone lib="adult ovary normalized (TLL)"
/note="Organ: pooled ovary; Vector: pBluescript II SK;
Site1: EcoRI; Site2: XhoI; Full length cDNA was
synthesized from adult ovary poly(A)+ RNA with ZAP-cDNA
Synthesis Kit (Stratagene). Size-fractionated cDNA (mainly
0.5-3kb in length) were directionally cloned into Uni-ZAP
XR vector. The cDNA clones were excised from Uni-ZAP
vector using ExAssist helper phage. Excised pBluescript
phagemids were used to transform E coli XL1-Blue cells.
The plasmid library was normalized by using a
reassociation-kinetics-based approach according to Bonaldo
et al., (Genome Research 6: 791-806,1996) with minor
modifications. The insert from randomly selected white
colonies was PCR amplified using M13 forward and reverse
primers and partially sequenced by using M13 reverse
primer."

```

ORIGIN

```

Query Match          17.0%; Score 322.4; DB 7; Length 640;
Best Local Similarity 69.6%; Pred. No. 4.4e-62;
Matches 437; Conservative 0; Mismatches 191; Indels 0; Gaps 0;

QY 851 TAGCCTCAGTCGATGCCCTTGACCTTTGGTCAGTCTCTGAGTACGAGAGCCCGCTGGGAGA 910
DB 13 TAGCTTTAGTTGATGCCCGACATTTGGAGAGTTTGGCCCTGCATGGCGCTCAGTGGGAAA 72

QY 911 CGGGCTAGGCTGCGGGGCACTGGCTTCCTGGCAGTACTTGGGTGGAGGATCGGTGC 970
DB 73 TGGGTGCAGGTTGCGGAGCAGTAGGCTTTCTTTCAGTGTCTGGCCCTCAGAGGGCTCATAC 132

QY 971 TGCTGCTCACTCTGCGCGCAGTGCAGTGCAGGCTCTCGTCTCTCTGTCGCGGGCTATG 1030
DB 133 TCTTCTGACTTTGGCCGAGTGCAGTGCAGTGTCTGTCTCATGTGCCCGCGCTTACG 192

QY 1031 GGAAGTCCCTCTCCCTGGGCGAGCTTCGAGCAGGGGTCTTAGGCTGCCTGGCAGTCGGCAG 1090
DB 193 GAAAGTACCACATCTTTGGGAAGTCCGAGCGCGTGCAGTTGCATGCTTGGCAGCTTTCGC 252

QY 1091 GGCTGGCGCGCCACTGCCCCCTGGCCCTCAGTGGGAGATACGGGGGCTCCCGCAGCTGCGC 1150
DB 253 TTGCAGTCGCTGCTTTGGCCACTAATTGGGGTGGAGAGTACGGGGCAACTCCAGCTCTGCA 312

QY 1151 TGCCCTACGCGCCACTGAGGCTCAGCCAGCAGCCCTGGGCTTACCGTGGCCCTGGTGA 1210
DB 313 TGCCATACCCCTACAGATGGAGAACCTTCCACACTGGGATTCATGGTTGCTCTCATAA 372

QY 1211 TGATGAACCTCCTTCTGTTTCTGTCGTGGCGGTGCTTACATCAAACTGTACTGTGACC 1270
DB 373 TGATGAACCTCTTTGTTGTTTCTGTTGATCACTGGAACTTACATCAAGTTGTACTGGGACC 432

QY 1271 TGCCCGGGGCGACTTTGAGGCGCGTGGGACTGCGCCATGCTGAGGACAGTGGCGCTGGC 1330
DB 433 TGATGAAGGGGCACTTTGACAGCAATTTGGGATTTGCCATGATTAAAGCATGTGGCGCTGGC 492

QY 1331 TCATCTTCGAGACGGGCTCTCTACTGTCCGCTGGGCTTCTCAGCTTCGCTCCATGC 1390
DB 493 TCATCTTCACCAACTGCTCTACTGCCCCGTTGCGCTTCCCTACCTTCCTCCCTCCCTTC 552

QY 1391 TGGGCTCTTTCCTGTCAGCGCCGAGGCGGTCAAGTCTGTCTGCTGGTGGTGTGCTGCCCC 1450
DB 553 TAAGGCTCTTTCGGTAAAGTGAAGAGGTTGTCAAACTGTGCTCCTAGTGTCTTTGCTC 612

QY 1451 TGCCTGCTGCTCAACCCACTGTGTGA 1478
DB 613 TTCCAGCAAGCATCAACCCCTTGTGTGA 640

```

Search completed: April 14, 2005, 18:37:13
Job time : 4102.82 secs

This Page Blank (uspto)